


```

source
1.1627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4475960"
/tissue_type="prostate, adenocarcinoma."
/clone_id="MH_MCC_91"
/lab_host="DH10B"
/Note="Vector: PCMV-SPORT6"
BASE COUNT      479 a      321 c      401 g      426 t
ORIGIN

Query Match      28.0% Score 506; DB 11; Length 1627;
Best Local Similarity 63.3%; Pred. No. 2e-117;
Matches 776; Conservative 0; Mismatches 450; Indels 0; Gaps 0;

QY 138 gylaatgatactgagcgaattcgtcgtcgcgcgaaccgtagtggtagaaggaaagt 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 GGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 198 ttgtatctatgacctgagtcgagcgaattactactatgacgaagtcacatccatgaa 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 CTGGTACTACTAGACGGGTGATATTGGAATTTATTATGAGACGGGTGATCCCATGAA 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 258 gcccacatgcacgcgaatgacccaatgcctcctcgcctcgaacgctcctcaagcatat 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 GCCCTCATAGATACCGCATTCGCTTAACCTGTTAAATTTATGCTTATACAGAAAAAT 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 318 gcaagttcgaagcctccctccgcgcgaagcgtgacatctgcgcgttcgaagcgaagaa 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 GGAATTTATAGGCCCATTAAGCCACTGCCGAAGAAATACAAATATTATCAGTGATGA 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 378 ctatgctctctctccgcagcacttaaccctgaaccgaacgaatcgaatcgcgaact 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 GTATATCAAAATTTCTACGGGCAATTAACACGATATGATGATAGTAAGCAGAT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 438 taagcgtcctcaagtgtagaagacgtccgcgtcttgaagcgtcttactcttgcga 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 540 GCAAGATTTAATGTTGGAGAAAGTTTCCAGTTGTTGATGGAATCTTTAGTTGTCTCA 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 498 gacataagtcggaagatctgttgtagctgtgcaagcttaaccagcgtcctcgcgaat 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 600 GCTCTCAACTGGCGGTTCACTGCTGAGCTGGAAGTTAAACCGACACAGACTGATAT 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 558 tgcacatcaactggcgtgtgtccatcaacgaatgaagtcgagcgtcctcgtctctg 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 GGCTGTTAATGGCTGAGAGATTAATCAATGTAAGAAATACAGAACATCAGGATCTCG 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 618 ttacgtcaatgatactgtcttagctactagagctccttaagcaagaatgaagcgtctct 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 720 TTACGTTAATGATATTGCTGCTTGGCAATCTTGAAATTAAGTATCATCAGAGAGCTT 779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 678 ttatgtcagatattgataccacacagagagtagtgaagagcatttatctactga 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 780 ATATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 738 caagggtatgactgtcgtttcaataaattggttacttcccggtacaggtacat 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 840 TCGGTATATGACGGTATCATCTCCATAATATGGGAATACCTTCCTGGACAGAGACTT 899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 798 tcaagatattgtagttagtcggaagaaatctactctcctcaagtagcactgtagatg 857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 900 GAGGGAATTTGGTCTGGAAAGCAAAATAGTATGCTGCTGAATTTTCCAAATGAGAGTGG 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 858 aatcgtatgtagagcactatcgttattccaagcccatcatgaggaattatagaat 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 960 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 918 ttccgagcagaggtgtgtgtagtcaagtgtgtgtagtccctactccttggagatg 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1020 GTATCAACGCTAGTGTGCTGTGATATACAGTGTGAGTCACTCAATATCTGTGTATGACT 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 978 aggtgtcctcaatcttcaatacaaggtcatgctgtagtgcgtcaaatattatgagatg 1037
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 1080 GGGTGTCTTCAATCTAACAGTCAMAGGTCATCTAATATGTTAGAGTTGTAACAACTTT 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1038 caatgtccctcactgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1140 TACCTTACCTTATGATCTGTTGAGAGAGTGTGTACAAATCCGTAATGTTGCTCGATG 1199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1098 ctgtgtcagagacgaggtgtgactgtgaggttgaagttgaagagaagaatgcgagaga 1157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1200 TTGGACATATGAGACTGACGATGCTGCTGATGTTGATGATGATGATGATGATGATGAT 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1158 tgaattatgaatcttgtgtccagactatataccttcaactgtgtgtccaaatgaatga 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1260 TGATTTCTTGTGATTTTGTGACACGACTTCAAACTGATATAGTCTTCAAAACATGAC 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1218 aataagaattctcgtcagatgcttgaagagattcgaatgactctccacatctc 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1320 AAGCCAGAACCTCCAGATATATGAGAAAGATAAACAGCTTGTGTAATTTGCG 1379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1278 taagcttcagacatgctccaaagtgtacatctcagaagaaccacatgatacagaactcc 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1380 CATGTACCTGATGACAGCTGTGTGTCAGATGACAGCTATTTCCAGAAAGATGTTGATGA 1439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1338 cgaagttgtagaagaccgaagatg 1363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1440 AGACAGTGTAGATGAAGATGAGAGAG 1465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
BH600970/c 774 bp DNA linear GSS 15-DEC-2001
LOCUS BH600970
DEFINITION BOGCN27TR BOGC Brassica oleracea genomic clone BOGCN27, DNA
sequence.
ACCESSION BH600970
VERSION BH600970.1 GI:17853416
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 774)
  Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
  Whole genome shotgun sequencing of Brassica oleracea
  Unpublished (2001)
  Other GSSs: BOGCN27TR
  CONTACT: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
    source
        1..774
            /organism="Brassica oleracea"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BOGCN27"
            /note="Vector: pPOS1; site_1: BstXI; 2-3 kb sheared
            genomic DNA inserted into pPOS1 using BstXI linkers"
BASE COUNT      222 a      194 c      174 g      184 t
ORIGIN

Query Match      26.8% Score 484.2; DB 12; Length 774;
Best Local Similarity 85.4%; Pred. No. 5.2e-112;
Matches 551; Conservative 0; Mismatches 93; Indels 1; Gaps 1;
QY 699 ccaaggagtagagtagagagcatttatgctactgacaggtatgactgtcgtt 758

```



```

Db 181 GTGTGAAGAGAAAGTTTATTCTATGACCCCTGAGCGCAATTAATACTATGCCC 240
    |||||
QY 242 aaggtatccatgaagcccatcgcataccgcatgaaccatgcctccctcgtcactaag 301
    |||||
Db 241 AAGGTATCCCAATGAAGCCCATCGATCGCATGCCATGCCCTCCGCTGCTAGC 300
    |||||
QY 302 gtctcttcagcatgaaggtctcaagccctcctcctccgcaagcgtatctgtcc 361
    |||||
Db 301 GTTCTTCAAGCATATGCAAGTCTCAAGCCCTTCTCCCTCCCGCAAGTATCTTGAC 360
    |||||
QY 362 gtcttcacgcg-acgacatgtctctctctcgcagcatcaccctgaaccacgaa 420
    |||||
Db 361 GNTTCCAGCGGACGACATATGCTCTCTTCTCCGAGCATTAACCTGAAACGACGAA 420
    |||||
QY 421 gatcagattcccaactaagcgtcgaatgtgtgaa 459
    |||||
Db 421 GATCAGATTGCG-CACCTTAAGCGGTNAATGTTGTGAA 458
    |||||

```

```

RESULT 8
LOCUS A1994124 457 bp mRNA linear EST 08-SEP-1999
DEFINITION 701499210 A. thaliana, Ohio State clone set Arabidopsis thaliana
ACCESSION A1994124
VERSION A1994124
KEYWORDS CDNA clone 701499210, mRNA sequence.
SOURCE EST
ORGANISM Arabidopsis thaliana

```

```

REFERENCE
AUTHORS Chen, J., Momiyama, M., Chan, F., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzozka, P.,
Gorjane, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
Rose, M., Warren, B., Ton, B., Kasbury, K., Borillo, C., Carpio, T.,
Policky, J., Suzuki, G., Argentine, C., Shah, S., Nohruga, A., Murry, L.,
Turner, C., Krikorian, S., Elder, L. and Hanson, D.
TITLE Arabidopsis thaliana Gene Expression Microarray
JOURNAL Unpublished (1999)
COMMENT Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

```

```

FEATURES
source
1.457
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="701499210"
/clone_lib="A. thaliana, Ohio State clone set"
/note="CDNA library was made from selected clones from the
Arabidopsis thaliana Ohio State clone set."
BASE COUNT 101 a 132 c 61 g 140 t 23 others
ORIGIN

```

```

Query Match 23.2% Score 418.8; DB 9; Length 457;
Best Local Similarity 94.1%; Pred. No. 1.7e-95;
Matches 432; Conservative 0; Mismatches 25; Indels 2; Gaps 1;
QY 1336 ccgaggtgatgaagaccagaagaatgggataaagaatggatcgatccagatcagacatg 1395
    |||||
Db 457 CCCGAGGTGATGAGACCAAGATGGGATGATGATGATGATGATGATGATGATGATG 398
    |||||
QY 1396 gatgtgatgatgacccgtataaccatacagaagagtaaaagaagaagcgtgttgaccca 1455
    |||||
Db 397 GATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 338
    |||||

```

```

QY 1456 gatcaaaagcaagatgactgtaagaagaattatggaagcgtggaagaaagtgtgagtg 1515
    |||||
Db 337 GATACTAAGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 278
    |||||
QY 1516 gaggtgatgagatggaagacacactaaggttacaggaaglaaacccagtgagtgagaa 1575
    |||||
Db 277 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 218
    |||||
QY 1576 gcaagtgtgaaataatggaagaggaagaaacaaacaaaggtgggagagcagcgttctc 1635
    |||||
Db 217 GCACGTGTGAAATGGAAGACGACGAAACAAACAAAGGGGCGGAGCAGCGCTTCTCT 158
    |||||
QY 1636 ccttaaacataagactcagagcttcaattcttctcacttcttctgtctatcaaatgt 1695
    |||||
Db 157 CCTAAACCTAAGACTCGAGCTTCTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 98
    |||||
QY 1696 gctagttaagttcttgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1755
    |||||
Db 97 GCTAGTTAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 38
    |||||
QY 1756 cagcagatgatactatcgttgcgtgcatgtctgaagatga 1794
    |||||
Db 37 CACGG--ATGCTTATTCTGTGATGCTGTGAATGATGA 1
    |||||

```

```

RESULT 9
LOCUS A1518877 999 bp mRNA linear EST 13-FEB-2001
DEFINITION A1518877 LTI_NFL011_NBC1 Homo sapiens CDNA clone CS0DA011YG11 5
ACCESSION A1518877
VERSION A1518877
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

```

FEATURES
source
1..999
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA011YG11"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA fax : (1) 301 610 8371 Email :
liang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 257 a 223 c 275 g 234 t 10 others
ORIGIN

```

```

Query Match 22.8% Score 412.6; DB 9; Length 999;
Best Local Similarity 66.0%; Pred. No. 8.9e-94;
Matches 620; Conservative 4; Mismatches 313; Indels 2; Gaps 2;

```



```

Db 245 TCGCCACTAGGCTCTCTGACGAGATGCTCAAGCCGACCCGCGCCGACCC 304
QY 350 gtgctctcgcgcgtctccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 409
Db 305 GCGACCTCTGCGCTTCCACGCGCAGCACTACCTCTCTCTCTCTCTCTCTCT 364
QY 410 aaaccgacgaatcatcatctccacacacacacacacacacacacacacacac 469
Db 365 AGAGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 424
QY 470 tcttgaagccttattccttcttgcacacacacacacacacacacacacacac 529
Db 425 TCTTCGACGCGCTCTACAGCTTCTGACGACGACGACGACGACGACGACGACG 484
QY 530 tcaagctaacacacacacacacacacacacacacacacacacacacacacac 589
Db 485 TCAAGCTCAACACGACGACGACGACGACGACGACGACGACGACGACGACG 541
QY 590 ctaagaagctgaagcctctgctctctctctctctctctctctctctctctct 649
Db 542 CCAAGAGAGTGGAGGCTCTCGGCTTCTGCTACGACGACGACGACGACGACG 601
QY 650 agctccttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 709
Db 602 AGCTCTCAATATACACGACGACGACGACGACGACGACGACGACGACGACGAC 661
QY 710 gagtgagagagagagagagagagagagagagagagagagagagagagagag 769
Db 662 GCGTGGAGGAGGCTTTTACACGACGACGACGACGACGACGACGACGACGAC 721
QY 770 gtgattacttccgctacagcagcagcagcagcagcagcagcagcagcagc 809
Db 722 GGGATTATTTTCCAGGACGACGACGACGACGACGACGACGACGACGACGAC 761

```

```

RESULT 11
BE053804 820 bp mRNA linear EST 07-MAR-2001
LOCUS BE053804
DEFINITION GA_Ea0035C07f Gossypium arboreum 7-10 dpa fiber library Gossypium
arborescens cDNA clone Ga_Ea0035C07f, mRNA sequence.
ACCESSION BE053804
VERSION BE053804.2 GI:13246713
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotyledons:
Rosidae: eufrosids II: Malvales: Malvaceae: Gossypium.
1 (bases 1 to 820)
Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.G., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
On Jun 8, 2000 this sequence version replaced gi:8380860.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 818.
Location/Qualifiers
1..820
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0035C07f"
/location="Gossypium arboreum 7-10 dpa fiber library"

```

```

FEATURES
SOURCE
1..820
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0035C07f"
/location="Gossypium arboreum 7-10 dpa fiber library"

```

```

/lisse_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 257 a 169 c 190 g 204 t
ORIGIN

```

```

Query Match 22.2% Score 400.6; DB 9; Length 820;
Best Local Similarity 76.4% Pred. No. 9, 2e-91;
Matches 519; Conservative 0; Mismatches 154; Indels 6; Gaps 2;

```

```

QY 937 gtattcaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 996
Db 1 GTCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
QY 997 atcaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1056
Db 61 ATCAAGGTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 1057 ttgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1116
Db 121 TTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 1117 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1176
Db 181 GTTGACCTGCGAAGGAGTGTGACGACGACGACGACGACGACGACGACGACG 240
QY 1177 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1236
Db 241 GGTGCGGATTTATACCTTCACTTGTGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 1237 atgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1296
Db 301 TTACTTGAAGAAATACGTAATTAATTAATTAATTAATTAATTAATTAATTA 360
QY 1297 agtgcacattcaagaagaagaagaagaagaagaagaagaagaagaagaaga 1356
Db 361 AGTGTCGAATTTTCAAGAAAGAACCACTGACGACGACGACGACGACGACGAC 420
QY 1357 gaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1416
Db 421 GATGATGACGATGAAGATGTGACCGGATTCAGACATGAGGTTGATGAGAGCG 477
QY 1417 cctatacgaagcagagtaaaagaagaagcgtgtgtgtgtgtgtgtgtgtgt 1476
Db 478 CCAATCCCTGTGACGAGTAAGAGAGAGCAATTAAGAACCGAGGAGGCT 537
QY 1477 ctgaagaagatttgaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1533
Db 538 CCGAAGGACCTGCAAGGACGACGACGACGACGACGACGACGACGACGACG 597
QY 1534 agcactaagttcagagagtaaaagcagtgagtgagtgagtgagtgagtgag 1593
Db 598 AGTACAAAGTCTGTAGATGTGACGCGCTATGATGATGATGATGATGATGAT 657
QY 1594 gaggaaggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1612
Db 658 CAAGAAACATCGAACAAAG 676

```

```

RESULT 12
AL531607 919 bp mRNA linear EST 13-FEB-2001
LOCUS AL531607
DEFINITION AL531607 L71_NFL001_NBC4 Homo sapiens cDNA clone CS00D0002YH15 5
prime, mRNA sequence.
ACCESSION AL531607
VERSION AL531607.1 GI:12795100
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```


Db 241 TTGAACCTTCACAAATACATGACGCGTGTGTGATGACATGATATCCACCATGCT 300
 QY 706 gatgaatgagagagacattatgctactacagaggtatagctgctcttccataaa 765
 Db 301 GATGCTGAGAGAGAGGCTTTTACACCAACGATAGGCTGATGCTTTCTTCATTAAG 360
 QY 766 ttggtgacttctcccgatagacatcagatagatagatagatagatagatagatag 825
 Db 361 TTGGGATTTACTTCTCCGACAGAGATGATGCTGATGATGATGATGATGATGATGAT 420
 QY 826 tactattctccatagacacactgagatgagatgagatgagatgagatgagatgagat 885
 Db 421 TATATTACTAAAGTTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 886 tccagaccatcagagagacattatgagaaatctccagacagagagagagagagagag 945
 Db 481 TTTAAGCCATATATGAGGAGAGGATTTATGAGATTTTAAAGCCCTGCTGTTTATCA 540
 QY 946 tgggtgctgac 1001
 Db 541 TGTGCTGCTGACCTTTATCTGGGACAGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 596

RESULT 15
 AL558916 847 bp mRNA linear EST 16-FEB-2001
 LOCUS AL558916 LTI_NFL008.TC2 Homo sapiens cDNA clone CS0DJ007YP20 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL558916
 VERSION AL558916.1 GI:12903904
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 847)
 AUTHORS Li, H.B., Gruber, C., Jessup, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
 source
 Location/Qualifiers
 1..847
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DJ007YP20"
 /clone_id="LTI_NFL008.TC2"
 /sex="male"
 /tissue_type="T cells from T cell leukemia"
 /note="Vector: pCMVSPORT 6; Site, 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 217 a 195 c 224 g 207 t 4 others
 ORIGIN

Query Match 21.4% Score 386.4; DB 9; Length 847;
 Best Local Similarity 67.1% Pred. No. 3.8e-87;
 Matches 546; Conservative 1; Mismatches 267; Indels 0; Gaps 0;

QY 189 gaggaatgttattatcattacacccgaggtcgacacattactactacagagagagagagag 248
 Db 34 GAGCAAACTCTGTTACTGACGAGCGGATGTTGAAATTTACTATTATGACAAAGGCCA 93

QY 249 tcccatgaagcccatcgcacatccgcacatgacacacacacacacacacacacacacacac 308
 Db 94 CCAATGAACCTCCACCGAATCCGCAATGATCAATTAATTTCTGCTCAACTATGCTGTCA 153
 QY 309 tcaagatagcagagctcgaagccctccctgcccgcagagagatgacatctgcccctcca 368
 Db 154 CCGAAATATGAATATATGACCCCTCAAAAGGCAATGCTGAGGAGATGACCAAGTACA 213
 QY 369 cgcgcagacatagctctctctccagacatlaacccctgaagacacagaaatcagat 428
 Db 214 CAGCGATGATCAATTAATTTCTGCGCTCAATCCGTCAGATTAATGTCGAGTACAG 273
 QY 429 tgcgcac 488
 Db 274 CAAGCAATGCAAGATTAACAGTGGTAGGACTGCGAGTATTTGATGAGCTGTTTGA 333
 QY 489 ctlttgccagac 548
 Db 334 GTTCTGCACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 593
 QY 549 ctgcacatagcacaac 608
 Db 394 GACGACATCGCTGTGATTTGGCTGGGAGCTGACACATGCAAAAGATCCAGGCTATC 453
 QY 609 tggcttctgac 668
 Db 454 TGGCTGCTGTTAGCTCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
 QY 669 ggcgttcttcttctgac 728
 Db 514 GAGGCTGCTTACTGATTCACATTTGATGATGATGATGATGATGATGATGATGATGAT 573
 QY 729 tgcac 788
 Db 574 CACCAACGACCGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
 QY 789 aggtcactcagagatagagatagagatagagatagagatagagatagagatagagat 848
 Db 634 TGGGACCTACGGGATATCGGGCTGGCAAAAGCATATTATGCTGTTAACTACCCCT 693
 QY 849 ggaatagagatagatagatagatagatagatagatagatagatagatagatagatag 908
 Db 694 CCAAGACGAGATGATGACGATGCTATGAGGCCATTTTCAAGCCGCTATGTCAAAAG 753
 QY 909 tatgaatcttcgac 968
 Db 754 AATGAGATGTTCCASCTAGTGGGCTTACAGTGTGCTGAGCTGAGCTGAGCTGAGCTG 813
 QY 969 ggaatagatagatagatagatagatagatagatagatagatagatagatagatag 1002
 Db 814 GGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847

Search completed: April 28, 2002, 18:51:38
 Job time: 7628 sec

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	740.4	41.0	1943	4	US-09-282-305-7	Sequence 7, Appl
2	698.6	38.7	2019	4	US-09-282-305-5	Sequence 5, Appl
3	559.4	31.0	1826	4	US-09-282-305-1	Sequence 1, Appl
4	498.2	27.6	2111	1	US-08-528-255A-2	Sequence 2, Appl
5	498.2	27.6	2111	1	US-08-717-365-2	Sequence 2, Appl
6	468	25.9	1576	4	US-09-282-305-9	Sequence 9, Appl
7	384.2	21.3	1475	4	US-09-282-305-3	Sequence 3, Appl
8	202.6	11.2	841	4	US-08-998-416-317	Sequence 317, Appl
9	85.8	4.7	7218	1	US-08-232-463-14	Sequence 14, Appl
10	50.8	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
11	44.6	2.5	289	4	US-09-007-005-17	Sequence 17, Appl
12	44.6	2.5	289	4	US-09-244-796-17	Sequence 17, Appl
13	38.8	2.1	289	4	US-09-007-005-17	Sequence 17, Appl
14	38.8	2.1	289	4	US-09-244-796-17	Sequence 17, Appl
15	38.8	2.1	485	4	US-09-177-325-4	Sequence 4, Appl
16	38.8	2.1	485	4	US-09-411-812A-4	Sequence 4, Appl
17	38.8	2.1	485	4	US-09-550-113-4	Sequence 4, Appl
18	38.2	2.1	3095	6	5231168-1	Patent No. 5231168
19	38	2.1	2518	4	US-09-433-699-3	Sequence 3, Appl
20	37.8	2.1	966	2	US-08-767-388-2	Sequence 2, Appl
21	37.8	2.1	3211	1	US-08-574-959A-8	Sequence 8, Appl
22	37.8	2.1	3211	4	US-09-357-014-8	Sequence 8, Appl
23	37.8	2.1	3901	2	US-08-574-959A-6	Sequence 6, Appl
24	37.8	2.1	3901	4	US-09-357-014-6	Sequence 6, Appl
25	37.4	2.1	696	4	US-09-461-697-193	Sequence 193, Appl
26	37.4	2.1	699	4	US-09-461-697-191	Sequence 191, Appl
27	37.4	2.1	717	4	US-09-461-697-189	Sequence 189, Appl

C	34	37.2	2.1	2295	1	US-08-375-300-3	Sequence 159, App1
	35	37.7	2.0	2295	1	US-09-177-431-3	Sequence 3, App11
	36	37	2.0	2295	5	PCT-US95-16930-3	Sequence 3, App11
	37	37	2.0	4080	3	US-08-375-300-1	Sequence 1, App11
	38	37	2.0	4080	3	US-09-177-431-1	Sequence 1, App11
	39	37	2.0	4080	5	PCT-US95-16930-1	Sequence 1, App11
	40	36.4	2.0	282	4	US-09-461-697-205	Sequence 203, App
	41	36.4	2.0	306	4	US-09-461-697-203	Sequence 1, App11
	42	36.4	2.0	2255	2	US-08-741-134-1	Sequence 1, App11
	43	36.4	2.0	9636	1	US-08-323-170B-1	Sequence 1, App11
	44	36.4	2.0	9636	4	US-08-954-441-1	Sequence 1, App11
	45	35.8	2.0	1505	1	US-07-315-246-1	Sequence 1, App11
	28	37.4	2.1	774	4	US-09-461-697-187	Sequence 187, App
	29	37.4	2.1	819	4	US-09-461-697-185	Sequence 185, App
	30	37.4	2.1	1669	3	US-09-461-697-184	Sequence 184, App
C	31	37.4	2.1	5356	4	US-08-658-136-2	Sequence 2, App11
	32	37.4	2.1	5357	3	US-08-658-136-1	Sequence 1, App11
	33	37.2	2.1	1934	3	US-08-961-083-159	Sequence 159, App
	34	37	2.0	2295	1	US-08-375-300-3	Sequence 3, App11
	35	37.7	2.0	2295	1	US-09-177-431-3	Sequence 3, App11
	36	37	2.0	2295	5	PCT-US95-16930-3	Sequence 3, App11
	37	37	2.0	4080	3	US-08-375-300-1	Sequence 1, App11
	38	37	2.0	4080	3	US-09-177-431-1	Sequence 1, App11
	39	37	2.0	4080	5	PCT-US95-16930-1	Sequence 1, App11
	40	36.4	2.0	282	4	US-09-461-697-205	Sequence 203, App
	41	36.4	2.0	306	4	US-09-461-697-203	Sequence 1, App11
	42	36.4	2.0	2255	2	US-08-741-134-1	Sequence 1, App11
	43	36.4	2.0	9636	1	US-08-323-170B-1	Sequence 1, App11
	44	36.4	2.0	9636	4	US-08-954-441-1	Sequence 1, App11
	45	35.8	2.0	1505	1	US-07-315-246-1	Sequence 1, App11

ALIGNMENTS

```

RESULT 1
US-09-282-305-7
Sequence 7, Application US/09282305
Patent No. 6287843
GENERAL INFORMATION:
APPLICANT: Baldwin, Donald A.
APPLICANT: Briggs, Steven P.
APPLICANT: Crane, Virginia C.
TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
FILE REFERENCE: 5718-44,
CURRENT APPLICATION NUMBER: US/09/282,305
CURRENT FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/080,563
PRIOR FILING DATE: 1998-04-03
NUMBER OF SEQ. ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1943
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (57)..(1610)
US-09-282-305-7

```

Query Match	41.08;	Score 740.4;	DB 4;	Length 1943;
Best Local Similarity	75.08;	Pred. No. 3e-230;		
Matches 953;	Conservative	0;	Mismatches 311;	Indels 6; Gaps 2

[illegible]


```

: APPLICANT: Baldwin, Donald A.
: APPLICANT: Briggs, Steven P.
: APPLICANT: Crane, Virginia C.
: TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
: FILE REFERENCE: 5718-44.
: CURRENT APPLICATION NUMBER: US/09/282,305
: CURRENT FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: 60/080,563
: PRIOR FILING DATE: 1998-04-03
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 9
: LENGTH: 1576
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (38)..(1336)
US-09-282-305-9

```

Query Match 25.9%: Score 468; DB 4; Length 1576;

Best Local Similarity 63.1%: Pred. No. 9.9e-142; Mismatches 430; Indels 3; Gaps 1;

Matches 739; Conservative 0; Mismatches 430; Indels 3; Gaps 1;

```

QY 217 gaagtcgcaactactactatgagcgaagatccatgagcccatgcatgcatg 276
  || || || || || || || || || || || || || || || || || || || ||
DB 83 gatgtgcaatgctactactgttgcccaatcccatgagcccatgcatgcatg 142
QY 277 accatgccccctccctgctactatgctctctttagatattgctgctcgaagcttc 336
  || || || || || || || || || || || || || || || || || || || ||
DB 143 acatcaacctgttcttcaatgacttcaataagaagatagatatataagccacac 202
QY 337 cctgcgcgcaagctgactctgctccgcttccacgcgcaagcactgctcttccgcg 396
  || || || || || || || || || || || || || || || || || || || ||
DB 203 aaagcatatccatagatagctgctcccatctcatctctgctattatgagaaattcttgac 262
QY 397 agcatcaccctgaagcccaagcagatcagattcgcacaaatgaagccttcaatgttgc 456
  || || || || || || || || || || || || || || || || || || || ||
DB 263 cgaataactcctgattccacagcactatgcaagtgaaactaataagatacaatcttga 322
QY 457 gaagacgctccgctcttgaagcgccttactctcttccgagcacttctgctgagagctc 516
  || || || || || || || || || || || || || || || || || || || ||
DB 323 gaagacgctccgctcttgaagcgccttactctcttccgagcacttctgctgagagctc 382
QY 517 gttgtgctctgctgaagccttaacacgagcctgctgcatattgccaactgagcgtgc 576
  || || || || || || || || || || || || || || || || || || || ||
DB 383 ttgagtctgctgagcagattcaatcaataatgctgacattgccaactgagcgtgc 442
QY 577 ggtctccatcaagcctgaagatgctgagcgtctgctgcttctgctgaatgatatctgc 636
  || || || || || || || || || || || || || || || || || || || ||
DB 443 gggctacatcagcacaagaagtgtgagcgtctgagcgtcttcttcaatgatatagta 502
QY 637 ttacgtacccagagcctccttaagcaagatgagcgttcttcttctgcatattgacatc 696
  || || || || || || || || || || || || || || || || || || || ||
DB 503 ttgagaaatccgagccttcttcaatgacatgcaaggggtcttcttcaatgacatgacg 562
QY 697 caccacgggagatgagatgagagcatttatgctactgacagggatgatactgctgc 756
  || || || || || || || || || || || || || || || || || || || ||
DB 563 catcatggagatgagatgagagccttcttcttcttctgacagggatgatactgctgc 622
QY 757 ttcatataattgtgta---ttacttcccgagacaggtccacatcagagatatggtat 813
  || || || || || || || || || || || || || || || || || || || ||
DB 623 ttccacaagatgagtgagcgttcttcttcttctgacagggatgatactgagagaa 682
QY 814 ggtgagcgaagatgatacttcttcaatgacatgacatgagatgagatgagagc 873
  || || || || || || || || || || || || || || || || || || || ||
DB 683 agggaaagaaataatgatacctacacatccacttaagatgagatgagagaaatagc 742
QY 874 tatcatcgttatacgaagcccatcatgaggaagtattgagaaattcttccagcagggct 933
  || || || || || || || || || || || || || || || || || || || ||
DB 743 ttctctgagccttcttcaacaataatgacgaaggtgtgtgagacatatctgctgctgc 802

```

```

QY 934 ggtgattgcaatgctgctgacccctatctgagatcgttgaattgttcaatctt 993
  || || || || || || || || || || || || || || || || || || || ||
DB 803 atgttctcaatgctgctgctgacccctatctgagatcgttgaattgttcaatctt 862
QY 994 tcatataaagatcgtgctgctgacccctatctgagatcgttgaattgttcaatctt 1053
  || || || || || || || || || || || || || || || || || || || ||
DB 863 tcatataaagatcgtgctgctgacccctatctgagatcgttgaattgttcaatctt 922
QY 1054 ctctggtgctgctgctgacccctatctgagatcgttgaattgttcaatctt 1113
  || || || || || || || || || || || || || || || || || || || ||
DB 923 gtaacggagaggtgctgacccctatctgagatcgttgaattgttcaatctt 982
QY 1114 ggaattgacatgagatgagatgagacgaagatgagcagatgagatgagatg 1173
  || || || || || || || || || || || || || || || || || || || ||
DB 983 ggggtcctttagacacgaagcctcccaatgagatgagatgagatgagatg 1042
QY 1174 ttggtccgacatcactcactgctgctgacccctatctgagatcgttgaattgtt 1233
  || || || || || || || || || || || || || || || || || || || ||
DB 1043 ttggtccgacatcactcactgctgctgacccctatctgagatcgttgaattgtt 1102
QY 1234 cagatgctgagagatgctgacccctatctgagatcgttgaattgttgaattgtt 1293
  || || || || || || || || || || || || || || || || || || || ||
DB 1103 acctatcctgctcactcactcactcactcactcactcactcactcactcactc 1162
QY 1294 ccaagtgtaccattcaggaagacacacacacacacacacacacacacacacac 1353
  || || || || || || || || || || || || || || || || || || || ||
DB 1163 cctggtgtcctcactcactcactcactcactcactcactcactcactcactc 1222
QY 1354 caagaagatgagatgagatgagatgagatgagatgagatgagatgagatg 1385
  || || || || || || || || || || || || || || || || || || || ||
DB 1223 gaatgagatcctgagatgagatgagatgagatgagatgagatgagatgagatg 1254

```

RESULT 7

US-09-282-305-3

; Sequence 3; Application US/09282305

; Patent No. 6287843

; GENERAL INFORMATION:

; APPLICANT: Baldwin, Donald A.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Crane, Virginia C.

; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses

; FILE REFERENCE: 5718-44.

; CURRENT APPLICATION NUMBER: US/09/282,305

; CURRENT FILING DATE: 1999-03-31

; PRIOR APPLICATION NUMBER: 60/080,563

; PRIOR FILING DATE: 1998-04-03

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Zea mays

; NAME/KEY: CDS

; LOCATION: (29)..(1084)

US-09-282-305-3

Query Match 21.3%: Score 384.2; DB 4; Length 1475;

Best Local Similarity 68.6%: Pred. No. 1.7e-114; Mismatches 243; Indels 0; Gaps 0;

Matches 530; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

```

QY 580 ctccatcagcagatgagatgagatgagatgagatgagatgagatgagatgagatg 639
  || || || || || || || || || || || || || || || || || || || ||
DB 161 cccacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 220
QY 640 gctatcctagatccttcttcttcttcttcttcttcttcttcttcttcttctt 699
  || || || || || || || || || || || || || || || || || || || ||
DB 221 gcatcctcagatccttcttcttcttcttcttcttcttcttcttcttcttctt 280
QY 700 cagcggagatgagatgagatgagatgagatgagatgagatgagatgagatg 759
  || || || || || || || || || || || || || || || || || || || ||

```



```
Db 1236 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1295
QY 451 gttggaagactgtccctcttgagcgccattatccttgcagcactgtgtga 510
Db 1296 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1355
QY 511 ggaactgtgtgtgtctctcaagcttaacacgagcctctgcgatatgtccatcaactg 570
Db 1356 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1415
QY 571 gctggtgtgtcctcactcaagcactaagatgc 600
Db 1416 YYYYYYYYYYYYYYYYYYGTACCAAAATTC 1445
```

RESULT 11

```
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17
```

```
Query Match 2.5%; Score 44.6; DB 4; Length 289;
Best Local Similarity 7.5%; Pred. No. 0.00014;
Matches 20; Conservative 114; Mismatches 131; Indels 0; Gaps 0;
```

```
QY 1385 attcagacatgagtgatgtgacgttaacctataccaagcaagataaagaagaag 1444
Db 15 ruruarcrurururururururururururururururururururururururur 74
QY 1445 ctgttgaccagatacaagaagacagatgtgactggaagaattatgagcggtggaaga 1504
Db 75 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 134
QY 1505 gttgtgaggtgagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1564
Db 135 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 194
QY 1565 gagtgaggaagcaagtggttaaaatggaagaaggaagcaagagtgagcgagc 1624
Db 195 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 254
QY 1625 aggcgtttcctctaacaataaga 1649
Db 255 rgrcruaaaaaaaaaaaaaaaaaa 279
```

RESULT 12

```
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17
```

```
Query Match 2.5%; Score 44.6; DB 4; Length 289;
Best Local Similarity 7.5%; Pred. No. 0.00014;
Matches 20; Conservative 114; Mismatches 131; Indels 0; Gaps 0;
```

```
QY 1385 attcagacatgagtgatgtgacgttaacctataccaagcagagataaagaagaag 1444
Db 15 ruruarcrurururururururururururururururururururururururur 74
QY 1445 ctgttgaccagatacaagaagacagatgtgactggaagaattatgagcggtggaaga 1504
Db 75 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 134
QY 1505 gttgtgaggtgagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1564
Db 135 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 194
QY 1565 gagtgaggaagcaagtggttaaaatggaagaaggaagcaagagtgagcgagc 1624
Db 195 rnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 254
QY 1625 aggcgtttcctctaacaataaga 1649
Db 255 rgrcruaaaaaaaaaaaaaaaaaa 279
```

RESULT 13

```
US-09-007-005-17/C
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 18:29:16 ; Search time 203.64 Seconds

(Without alignments)
15235.044 Million cell updates/sec

Title: US-09-645-337-1

Perfect score: 1807
Sequence: 1 agagagcagctccctccccc.....atgatgatgatgatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_032802:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1807	100.0	1807	22	AAE80350
2	830.2	45.9	1805	21	AAZ58260
3	740.4	41.0	1943	20	AAZ580840
4	734.2	40.6	1990	21	AAZ58259
5	698.6	38.7	2019	20	AAZ580839
6	559.4	31.0	1826	20	AAZ580837
7	525	29.1	1609	21	AAZ580837
8	525	29.1	1611	21	AAZ580837
9	523.4	29.0	1800	22	AAE80351

10	508.8	28.2	1997	24	AB19512
11	504.4	27.9	1985	20	AAZ32066
12	504.4	27.9	1985	22	AAZ32066
13	504.4	27.9	1985	22	AAZ32066
14	499.8	27.7	1449	18	AAZ58371
15	499.8	27.7	1611	22	AAZ58371
16	499.8	27.7	2163	21	AAZ58371
17	498.2	27.6	2111	17	AAZ58371
18	495.8	27.4	2145	23	AB101897
19	468	25.9	1576	20	AAZ58371
20	451.8	25.0	1954	22	AAZ58371
21	421.4	23.3	4755	23	AAZ58371
22	414.8	23.0	1582	23	AAZ58371
23	384.2	21.3	1475	20	AAZ58371
24	381.8	21.1	1662	23	AB104121
25	330.2	18.3	3793	23	AB104121
26	311.8	17.3	2009	22	AB104121
27	255.6	14.1	3826	23	AB104121
28	212.6	11.8	1682	22	AAZ58371
29	211	11.7	1367	22	AAZ58371
30	211	11.7	1654	21	AAZ58371
31	211	11.7	2046	22	AAZ58371
32	186	10.3	541	21	AAZ58371
33	170.4	9.4	402	22	AAZ58371
34	166.6	9.2	379	18	AAZ58371
35	150.2	8.3	375	18	AAZ58371
36	143	7.9	1539	22	AAZ58371
37	134.6	7.4	1016	22	AAZ58371
38	131.6	7.3	570	21	AAZ58371
39	121.8	6.7	241	22	AAZ58371
40	121.8	6.7	241	22	AAZ58371
41	105.6	5.8	227	18	AAZ58371
42	99.6	5.5	222	22	AAZ58371
43	97.6	5.4	411	22	AAZ58371
44	96.4	5.3	286	23	AAZ58371
45	89.4	4.9	399	22	AAZ58371

ALIGNMENTS

RESULT 1	AAE80350	AAE80350 standard; DNA: 1807 BP.
ID	AAE80350;	
AC	AAE80350;	
XX		
DT	29-JUN-2001 (first entry)	
XX		
DE	Nucleotide sequence of a histone deacetylase designated ATRPD3A.	
XX		
KW	Histone deacetylase: ATRPD3A; RPD3; gene expression; transgenic plant;	
KW	HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.	
XX		
OS	Arabidopsis thaliana.	
XX		
FH	Key	Location/Qualifiers
FT	142..1647	/*tag- a
FT	CDS	/product- "histone deacetylase ATRPD3A"
XX		
PN	CA2316036-A1.	
PD	27-FEB-2001.	
XX		
PF	24-AUG-2000; 2000CA-2316036.	
XX		
PR	27-AUG-1999; 99US-0383971.	
XX		
PA	(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.	
XX		
PI	Miki B, Brown D, Tian L, Wu K;	
XX		

Mouse ischaemic co
Human MERR2 relate
U131814 cDNA clone.
Human histone deac
cDNA encoding a hi
Human histone deac
Human prostate can
RPL1 transcritpion
Drosophila melanog
Human histone deac
Human histone deac
Drosophila melanog
DNA encoding novel
Maize histone deac
Drosophila melanog
Drosophila melanog
Human bone marrow
Drosophila melanog
Nucleotide sequenc
Human histone deac
Human polynucleoti
Wheat histone deac
Human breast tumor
Partial cDNA encod
Human polynucleoti
cDNA encoding nove
Fusarium venenatum
Human breast cance
Human breast cance
Partial cDNA encod
Human ovavian can
Novel human diagno
cDNA #417 encoding
Novel human polynu


```

Db 1661 tgcatacaaatgctgttaagattcttcggagtgctgtgtgtgtgtgaagaccctccctg 174
Oy 1741 tttaagagagattgcacgagatattgattatcgttcgacgtctgaatgatgatgatg 1800
Db 1741 tttaagagattgcacgagatattgattatcgttcgacgtctgaatgatgatgatg 1800
Oy 1801 atgacaa 1807
Db 1801 atgacaa 1807

RESULT 2
AA258260
ID AA258260 standard; cDNA; 1805 BP.
AC AA258260;
CT 08-MAY-2000 (first entry)
DE Soybean histone deacetylase gene 1 (HD1) cDNA clone srl.pk0023.d1.
KW Chromatin associated protein; histone deacetylase gene 1; HD1;
   soybean; transgenic plant; transcription regulation; ss.
XX Glycine max.
FH Key location/Qualifiers
FT CDS 130..1551
FT     /tag= a
XX
XX MO200004177-A1.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99MO-US15807.
XX
XX 14-JUL-1998; 98US-0092841.
XX
XX (DUP0 ) DU FONT DE MEMOURS & CO E I.
XX
XX Cahoon RE, Vollmer SJ;
XX
XX MPI: 2000-182439/16.
XX DR P-PSDB: AAY58829.
XX
XX New nucleic acid fragment useful as probes and primers, for
XX transforming plants -
XX
XX Claim 3; Page 26-27; 36pp: English.
XX
XX The present sequence is that of cDNA clone srl.pk0023.d1 encoding
XX soybean histone deacetylase 1 (HD1, see AAY58829), a chromatin
XX associated protein. The cDNA clone was isolated from a soybean
XX root cDNA library on the basis of homology to other plant histone
XX deacetylases. The invention relates to isolated rice, soybean and
XX wheat nucleic acid fragments encoding HD1. It also relates to the
XX construction of a chimeric gene encoding all or a portion of HD1,
XX in sense or antisense orientation, where expression of the chimeric
XX gene results in production of altered levels of HD1 in a
XX transformed host cell. The availability of nucleic acid sequences
XX encoding (portions) of histone deacetylase proteins will facilitate
XX studies of global transcriptional regulation in eukaryotic cells,
XX and will also provide mechanisms to control transcriptional gene
XX regulation in plants.
XX
XX Sequence 1805 BP; 524 A; 309 C; 436 G; 534 T; 2 other:

Query Match 45.9%; Score 830.2; DB 21; Length 1805;
Best Local Similarity 76.1%; Pred. No. 4.4e-244;
Matches 1021; Conservative 1; Mismatches 319; Indels 0; Gaps

128 acgaaaaagaglaatgacactcgtgcgcgaatcgtcgtgcgtccgaccctgatgtatga 187

```

Db	116	acgaaacacgagaaacgaaacgagaggaacccctccacacagagctcagatgctgga	175
Qy	188	agaagaaagctgttatcttatccatgaccgcgaggtgcgcacattacactatgcacgaagtc	247
Db	176	agaagaaagcttatcatattcttatgacacgaaggttggaaactatattatgagcaggaac	235
Qy	248	atcccatgaagcccacgcgacccgcacatgacccatgcctcctcgtcactaagctcc	307
Db	236	accacaatgaaacacacacagatccgatgcacatgcctcttlaagcccaatgatgtgc	295
Qy	308	ctcagcatatgacaggttctcaagccctccctgcgcgcgaaagctatctcgcgccttc	367
Db	296	ctcaaccacatgcaggtctgaaacgactatgctgctaaagatagagaccttgcgaattcc	355
Qy	368	acgcgcgcgacatgctcctctctccgcgcacatlaaccccttgaacccgcgaacatcga	427
Db	356	atgctgatgatattatgctgcctctctcgaagagacacaccccttgaacgcgacgaagctac	415
Qy	428	ctgcacaaactaagcgcctcaatgctgtgtgaagaactgcgcgcttctgaacgcctta	487
Db	416	tgaagacagctgaagaggttcttaatgcttcggaagaactgcgccttatattgattgtct	475
Qy	488	cccttgcgcgaacctatgctcgcgcgcgcgcgcgttgcgttccttcaagcttaacccgcgc	547
Db	476	ctcttcgcgcgaacatcgcgcgcgcgcgcgcgtctgtgtgtcctcaagttgcgaacctgag	535
Qy	548	ctctgcgatattgcacatcaactgcgcgcgcgcgcgcctccacacgaagaatgcgcgcgc	607
Db	536	tatgtgatattgcacataaattgcgcgcgcgcgcgcctacacatagcgaagaagctgtgcgc	595
Qy	608	ctgcgcctctgtlaagctacatgatacgtctttagctactcctagaactccttgaacgcac	667
Db	596	ctgcgccttgcctatgttaatgacatcgtgcgtgcctatttggaaacttccaaatacgt	655
Qy	668	agcgtctcttcttatgctcgcataatgataccacacgcgcgcgcgcgcgcgcgcctt	727
Db	656	agcgtctctctgatagtgcacatgataccacacgcgcgcgcgcgcgcgcgcgcctt	715
Qy	728	atgctgcgcgaagcgcgtatgaactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	787
Db	716	acacacacgc	775
Qy	788	cagctgcacatcaggaatagatagttatgtagtcgcgcgcgcgcgcgcgcgcgcgcgc	847
Db	776	cagctgcacatcgc	835
Qy	848	tgc	907
Db	836	tgc	895
Qy	908	ctatgc	967
Db	896	ctatgc	955
Qy	968	gc	1027
Db	956	gc	1015
Qy	1028	tgc	1087
Db	1016	tgc	1075
Qy	1088	tgc	1147
Db	1076	tgc	1135
Qy	1148	tgc	1207
Db	1136	tgc	1195
Qy	1208	gtaacatggaataaagaactctcgttcgaagcttgcgaagaatgcgaatgcaccttcc	1267

Db	1091	gtactatgaatacttcgctgcgcagattaaacctcttcattgttcacaaagttaacatgaggaa	1150
Qy	1221	taagaattctctgcacaaatctcttgaagaagatttcgcaatgacctcttcacaattctcttaa	1280
Db	1151	caaaaataacccgcacaaacaaactggatagataaagatcctaactctcgataatcttcaaa	1210
Qy	1281	gcttcacagatctccacagtgataccallcagaagaagaccacctgatcaagaagactccga	1340
Db	1211	acttcgacatgcctccctcagtcgtccacttcaagaagagagcttcctgcacagaataactga	1270
Qy	1341	ggttcgttcgaagaccaaagaatgaggatgaataagaatggagatccgatatcaacatgatgt	1400
Db	1271	gcacagatgaagaatccaagaatgatcccgatgaagaagaatgatctctgaactctgtatagaaat	1330
Qy	1401	tgatgatgatgc	1410
Db	1331	ggatggaccac	1340

RESULT 4

ID AAZ58259 standard; cDNA; 1990 BP.

AC AAZ58259;

DT 08-MAY-2000 (first entry)

DE Rice histone deacetylase gene 1 (HD1) cDNA contig.

KW Chromatin associated protein; histone deacetylase gene 1; HD1;

rice; transgenic plant; transcription regulation; ss.

05 *Oryza sativa*.

FH	Key	Location/Qualifiers
33		33

FT	CDS	3..1484
----	-----	---------

/partial

PN WO200004177-A1.

PD 27-JAN-2000.

PF 13-JUL-1999; 99MO-US15807.

PR 14-JUL-1998; 98US-0092841.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Cafoon RE, vollmer SJ;

DR WPI: 2000-182439/16.

DR P-PSDB; AAY58828.

PT New nucleic acid fragment useful as probes and primers, for

PT transforming plants

PS Claim 3; Page 24; 36pp; English.

CC The present sequence is that of a cointg encoding a portion
CC (see AAY58828) of rice histone deacetylase 1 (HD1), a chromatin
CC associated protein. The cointg was obtained from cDNA clones
CC isolated from rice callus, 15-day-old test and 15-day-old seedling
CC cDNA libraries on the basis of homology to other plant histone
CC deacetylases. The invention relates to isolated rice, soybean and
CC wheat nucleic acid fragments encoding HD1. It also relates to the
CC construction of a chimeric gene encoding all or a portion of HD1,
CC in sense or antisense orientation, where expression of the chimeric
CC gene results in production of altered levels of HD1 in a
CC transformed host cell. The availability of nucleic acid sequences
CC encoding (portions) of histone deacetylase proteins will facilitate
CC studies of global transcriptional regulation in eukaryotic cells,

CC and will also provide mechanisms to control transcriptional gene
CC regulation in plants.

SQ Sequence 1990 BP; 502 A; 494 C; 518 G; 476 T; 0 other;

Query Match 40.68; Score 734.2; DB 21; Length 1990;

Best Local Similarity 76.0%; Pred. No. 1.5e-214;

Matches 920; Conservative 0; Mismatches 288; Indels 3; Gaps 1;

[illegible]


```

Db      1082 cccgattatactctcatalcccaacaaagtgtgaacacctgaataccacaagac 1141
Oy      1240 ctgaagagatctgcgaatgcaccttccacatctcttaagatcgaatgcgaat 1299
Db      1142 ttggaagaacataagacatgatattggaagaccctgcaaaagataacatgttccagc 1201
Oy      1300 gtaccattcagaagaacaccctgatcacagagactcccgaggttgatgaga 1352
Db      1202 acccaatccatgacagacgcgtcagaccctgaagctcgaagagagaagaaga 1254

RESULT 7
ID      AAC46043 standard; DNA: 1609 BP.
AC      AAC46043:
XX
XX      18-OCT-2000 (first entry)
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 48704.
DE
XX      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX
XX      Arabidopsis thaliana.
XX      EP1033405-A2.
XX
XX      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-0301439.
PF
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
XX      09-MAR-1999; 99US-0123548.
XX      23-MAR-1999; 99US-0125788.
XX      25-MAR-1999; 99US-0126264.
XX      29-MAR-1999; 99US-0126785.
XX      01-APR-1999; 99US-0127462.
XX      06-APR-1999; 99US-0128234.
XX      08-APR-1999; 99US-0128714.
XX      16-APR-1999; 99US-0129845.
XX      19-APR-1999; 99US-0130077.
XX      21-APR-1999; 99US-0130449.
XX      23-APR-1999; 99US-0130510.
XX      28-APR-1999; 99US-0130891.
XX      30-APR-1999; 99US-0131449.
XX      04-MAY-1999; 99US-0132407.
XX      05-MAY-1999; 99US-0132484.
XX      06-MAY-1999; 99US-0132485.
XX      07-MAY-1999; 99US-0132486.
XX      07-MAY-1999; 99US-0132487.
XX      11-MAY-1999; 99US-0134256.
XX      14-MAY-1999; 99US-0134218.
XX      14-MAY-1999; 99US-0134219.
XX      14-MAY-1999; 99US-0134221.
XX      14-MAY-1999; 99US-0134370.
XX      18-MAY-1999; 99US-0134768.
XX      19-MAY-1999; 99US-0134941.
XX      20-MAY-1999; 99US-0135124.
XX      21-MAY-1999; 99US-0135353.
XX      24-MAY-1999; 99US-0135629.
XX      25-MAY-1999; 99US-0136021.
XX      27-MAY-1999; 99US-0136392.
XX      28-MAY-1999; 99US-0136782.
XX      01-JUN-1999; 99US-0137222.
XX      03-JUN-1999; 99US-0137528.
XX      04-JUN-1999; 99US-0137502.
XX      07-JUN-1999; 99US-0137724.

PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145086.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145224.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
```


Db 1194 acacgcccacgaatagagagataagacacacgttgcgtgacacattcggactaa 1253
Oy 1286 agcatgctccaaatgtagacatttcaggaagaagaccacct 1324
1254 taacgacacacgttcacgttcacgttcacacacaccaccag 1292

RESULT 8

AAC33022
ID AAC33022 standard; DNA; 1611 BP.

AC AAC33022;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1498.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0123180.
XX 09-MAR-1999; 990S-0123548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 16-APR-1999; 990S-0128714.
XX 19-APR-1999; 990S-0129845.
XX 21-APR-1999; 990S-0130077.
XX 23-APR-1999; 990S-0130049.
XX 23-APR-1999; 990S-0130510.
XX 23-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 30-APR-1999; 990S-0132407.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 07-MAY-1999; 990S-0132487.
XX 11-MAY-1999; 990S-0132863.
XX 14-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134218.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 18-MAY-1999; 990S-0134370.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 24-MAY-1999; 990S-0135629.
XX 25-MAY-1999; 990S-0136021.
XX 27-MAY-1999; 990S-0136392.
XX 28-MAY-1999; 990S-0137282.
XX 01-JUN-1999; 990S-0137222.
XX 03-JUN-1999; 990S-0137528.
XX 04-JUN-1999; 990S-0137502.
XX 07-JUN-1999; 990S-0137724.
XX 08-JUN-1999; 990S-0138094.
XX 10-JUN-1999; 990S-0138540.
XX 10-JUN-1999; 990S-0138847.
XX 14-JUN-1999; 990S-0139119.

PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145277.
PR 27-JUL-1999; 990S-0145278.
PR 27-JUL-1999; 990S-0145279.
PR 27-JUL-1999; 990S-0145280.
PR 27-JUL-1999; 990S-0145281.
PR 28-JUL-1999; 990S-0145282.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0146389.
PR 04-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 05-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147392.
PR 05-AUG-1999; 990S-0147600.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.

PR 12-AUG-1999: 990S-0148341.
PR 13-AUG-1999: 990S-0148565.
PR 13-AUG-1999: 990S-0148684.
PR 16-AUG-1999: 990S-0149368.
PR 17-AUG-1999: 990S-0149175.
PR 18-AUG-1999: 990S-0149426.
PR 20-AUG-1999: 990S-0149722.
PR 20-AUG-1999: 990S-0149723.
PR 20-AUG-1999: 990S-0149929.
PR 23-AUG-1999: 990S-0149902.
PR 23-AUG-1999: 990S-0149930.
PR 25-AUG-1999: 990S-0150566.
PR 26-AUG-1999: 990S-0150884.
PR 27-AUG-1999: 990S-0151065.
PR 27-AUG-1999: 990S-0151080.
PR 30-AUG-1999: 990S-0151303.
PR 31-AUG-1999: 990S-0151438.
PR 01-SEP-1999: 990S-0151930.
PR 01-SEP-1999: 990S-0152363.
PR 10-SEP-1999: 990S-0153070.
PR 13-SEP-1999: 990S-0153758.
PR 15-SEP-1999: 990S-0154018.
PR 16-SEP-1999: 990S-0154039.
PR 20-SEP-1999: 990S-0154779.
PR 22-SEP-1999: 990S-0155139.
PR 23-SEP-1999: 990S-0155486.
PR 24-SEP-1999: 990S-0155659.
PR 28-SEP-1999: 990S-0156458.
PR 29-SEP-1999: 990S-0156596.
PR 04-OCT-1999: 990S-0157117.
PR 05-OCT-1999: 990S-0157753.
PR 06-OCT-1999: 990S-0157865.
PR 07-OCT-1999: 990S-0158029.
PR 08-OCT-1999: 990S-0158232.
PR 12-OCT-1999: 990S-0158369.
PR 13-OCT-1999: 990S-0159293.
PR 13-OCT-1999: 990S-0159294.
PR 13-OCT-1999: 990S-0159295.
PR 14-OCT-1999: 990S-0159329.
PR 14-OCT-1999: 990S-0159330.
PR 14-OCT-1999: 990S-0159331.
PR 14-OCT-1999: 990S-0159637.
PR 14-OCT-1999: 990S-0159638.
PR 18-OCT-1999: 990S-0159584.
PR 21-OCT-1999: 990S-0160741.
PR 21-OCT-1999: 990S-0160767.
PR 21-OCT-1999: 990S-0160768.
PR 21-OCT-1999: 990S-0160770.
PR 21-OCT-1999: 990S-0160814.
PR 21-OCT-1999: 990S-0160815.
PR 22-OCT-1999: 990S-0160980.
PR 22-OCT-1999: 990S-0160981.
PR 22-OCT-1999: 990S-0160989.
PR 25-OCT-1999: 990S-0161404.
PR 25-OCT-1999: 990S-0161405.
PR 25-OCT-1999: 990S-0161406.
PR 26-OCT-1999: 990S-0161359.
PR 26-OCT-1999: 990S-0161360.
PR 26-OCT-1999: 990S-0161361.
PR 28-OCT-1999: 990S-0161920.
PR 28-OCT-1999: 990S-0161992.
PR 28-OCT-1999: 990S-0161993.
PR 29-OCT-1999: 990S-0162142.

Query Match 29.1%; Score 525; DB 21; Length 1611;
Best Local Similarity 66.0%; Pred. No. 2.6e-150;
Matches 778; Conservative 0; Mismatches 395; Indels 6; Gaps 1;

OY 152 ggggaattgcgtgggtccggagcagatggtgaaagaaagtgttattctatg 211
DB 116 ggggaattgcgtgggtccggagcagatggtgaaagaaagtgttattctatg 175

OY 212 accctagctcgcaattactactatgccaagttacatccataagcccatccac 271
DB 176 agccgagcagctggagactactactacgttcaaggccaaccggtgaagctcaacgagcc 235
OY 272 gaatgaccatgcctctcgtctcaactacgtctccttcaagatalgtcagttccaacg 331
DB 236 gtaatgcatatagcctaactatccatccacacgctcgttgaatcagtcgc 295
OY 332 ccttcctcccgccgaacgtatctcgtccgttccacgcgaacgactatgctcttc 391
DB 296 ctgagcctcgtaacgcttcctacatccgcgaattccatccgcggagtatgttgcattc 355
OY 392 tccgcagcattaccccttgaaacccgcaagatc-----agttccgccaattagcct 445
DB 356 tgccttcgcttcgcgcggaatctatggcgatccttcgcgtcacaacccaagcgat 415
OY 446 tcaatgttgtaagactgtccgtctcttgaagccttattcctttgccaagactatg 505
DB 416 tcaatgttgtaagactgtccgtctcttgaagccttattcctttgccaagactatg 475
OY 506 ctggagagatcgttggctgtcgtcgaagcttaacacagcgctcgtgatalccacata 565
DB 476 ccggaagttctatgtgtgtcgtccgtaaatlaacagacaggaacgtatccgatalca 535
OY 566 actggagctgtgtctccacgcctaaagatgtcgaagcctcgtctgttlaacgta 625
DB 536 attggagcgtgtggcttccacatgtcctaaagaaacgaggtcttcggttgcctatgaa 595
OY 626 atgatatcgtcttaagctacccctagagctccttaagcaagatgagctgttcttatgct 685
DB 596 acgacatcgtcgtcgaagctcgtgagcttccgaagcgttcttaccacatgataagata 655
OY 686 atattgatccaccccgaggtgaggtgaggaagcatttactactgacaaggtta 745
DB 656 atattgatccaccccgaggtgaggtgaggaagcgttcttaccacatgataagata 715
OY 746 tgaactcgtcttcaataattgtgattacttcccggtcgaagctacatccagata 805
DB 716 tgaactcgtcttccacaatttgggacttcttcccaagaaactgtccalaagagatg 775
OY 806 taggtatgttagcgaaagatctatctcctcaatgttaccatgtgatgataatcgtatg 865
DB 776 ttggcgtctgaagaaaggaatactatgctcctaattgttccactaaacgtgtatgagcg 835
OY 866 atgagagcatcatctgttattcaagcccatatcgtgggaagattatgaaatttccagc 925
DB 836 atgaaagttccgcaagctgtttagaccccttataccagaagttatggaagttacagc 895
OY 926 caagggcgtgtgattgcaatgtgtgctgactcctcatcttgaggatcggttagttgct 985
DB 896 caagggcagttgtcttccagctgtgtgcgaaccccttaagtggtatcggtatggtgct 955
OY 986 tcaatcttcaatcaaaagtcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1045
DB 956 tcaactatcagtcgaaggtcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1015
OY 1046 ccttactgctcttgggtgtgtgtgttatactatccgcgaatgttccgctgtcgtgct 1105
DB 1016 ctctcatgt 1075
OY 1106 acgagactgagttgacttgcagttgagttgagttgagttgagttgagttgagttgagttg 1165
DB 1076 atgagactgagttgacttgcagttgagttgagttgagttgagttgagttgagttgagttg 1135
OY 1166 atgaaacttgcgtccagactatcacatcagttgctcgaatgaatgaatgaatgaatga 1225
DB 1136 atgaaacttgcgtccagactatcacatcagttgctcgaatgaatgaatgaatgaatga 1195
OY 1226 attctcgtcagatgcttgaagatgattgcgaatgacttctcacaatctcctaagcttc 1285
DB 1196 acagcccaaaatataggagagatgaagaaacagttgctgaaacaacttccggaactaa 1255
OY 1286 agcatgtccaaatgtaacatttccagaaagcaccactg 1324

XX AA32066;
AC
XX
DT 10-JAN-2000 (first entry)
XX
DE Human METH2 related EST U31814.
XX
KW Human; METH2; METH2; anti-angiogenic; metalloproteinase thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency; ss.
XX
OS Homo sapiens.
XX
PN M09937660-A1.
XX
PD 29-JUL-1999.
XX
PE 22-JAN-1999; 99MO-US01313.
XX
PR 23-JAN-1998; 98US-0072298.
PR 28-AUG-1998; 98US-0098539.
XX
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBR/) RUBEN S M.
XX
PI IrueLA-Arispe L, Hastings GA, Ruben SM;
DR WPI: 1999-590684/50.
XX
PT New isolated metalloproteinase thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders
PT -
XX
PS Disclosure; Page 444-446; 457pp; English.
XX
XX AA32000 and AA32001 encode, and AAY49501 and AAY49502 represent, human
XX metalloproteinase thrombospondin (METH) proteins METH1 and METH2
XX respectively. METH1 and METH2 have been found to be potent inhibitors of
XX angiogenesis both in vitro and in vivo. They can be used for treating
XX cancer and other disorders related to angiogenesis including abnormal
XX wound healing, inflammation, rheumatoid arthritis, psoriasis,
XX endometrial bleeding disorders, diabetic retinopathy, some forms of
XX macula degeneration, haemangiomas, and arterial-venous malformations.
XX They may be useful in treating deficiencies or disorders of the immune
XX system, by activating or inhibiting the proliferation, differentiation,
XX or mobilisation (chemotaxis) of immune cells. The etiology of these
XX immune deficiencies or disorders may be genetic, somatic, such as
XX cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
XX toxins), or infectious. They can also be used to treat inflammatory
XX conditions, both chronic and acute conditions. The products can also be
XX used for detection and diagnosis. AA32002 to AA32080, and AAY49503 to
XX AAY49511 represent sequences given in the exemplification of the present
XX invention.
XX
XX Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

QY	258	gcccatacgatcccgatgtgcccataccgcccctccgctccaaatcgctcccttcacgatat	317
Db	300	gccctacagaatcccgcatgagccataactgtcgttaaatcttgcttatcacagaanaat	359
QY	318	gaagcttccaaagccctccctccgcccggaagtgatctctgcgccttcacacgcgcga	377
Db	360	ggaatatatagagcccatlaaagccactgcccgaagaatgacaaatatcacagtgatga	419
QY	378	ctatgctctcttctccgcagcatctaacccctgaaccccgcaagatcagatccgcaact	437
Db	420	gtatatacaaatctctacggtacataaagaccagaataacatgctctgagtgatgaagcaagat	479
QY	438	taagcgcttcacatcttggtgaagactgcgtccgctcttggttcgcgccttatcttccttgca	497
Db	480	gcataatattaaatcgtcggagaagaagatgcccagcggttttgatgaaccttgatgttcga	539
QY	498	gaacctatctggaagaaactcttggtgctctgtccaaagcttcaacacacgcgctctgcgatat	557
Db	540	gctcccaactcggcggttcagttcgtgtggagcttggaagttcaacgcgaacacagacgtatga	599
QY	558	tgccaatcaactggcgctggtgctcccaatcagctcaagaagtgccgaagcctctgctctg	617
Db	600	ggcgtttaatctggcgaggtatcacatcatcgtctaaagaatccgaagcctcaggatctctg	659
QY	618	ttacgcataagatctgcgtcttagctatcctagagcgcccttaagcgcaagcagcggtctc	677
Db	660	ttacgtttaaagatctgctgtctgcgcaccccttggaattcaacaaagttacatcacagaagctct	719
QY	678	ttatgctcatatctatataccacacacgggagtgagtgagtgagggcgattttatgctactga	737
Db	720	atataatgatatagatatactatcatcatggttgagtggtgltggaagaagccttttatacaacaga	779
QY	738	cagggtatgacgctcctcggttcaataattgtgtatcctctcccggtacaggtccat	797
Db	780	tcgtgtaatgacgggtatcatcatcacaatacttgaggaaatcctcttcggacagaggaagctt	839
QY	798	tcagatatatggttatgttgtagcgaaagtgactatcttcacaaagtgaaacctgtgagtgat	857
Db	840	gagggataatctggtcgtggaagaagccaatactactgcgtccaattcttcaatctgctgagtg	899
QY	858	aatcagtatatgagcctatcatcctgttatccaagcccatcatcgtgggaagaagttatgaaat	917
Db	900	tatagatgatgagcgtatcatcagggacagatattgaagcctatattcacaaggtgagtgagat	959
QY	918	tttcgcgacgaaggcgctggtattgccaatggtgctgtagccctcatctcggagatcggtt	977
Db	960	gtatacaacctgagctgtggttatctaacagttggtgtagacatcatatctctggtgatagact	1019
QY	978	aggttgcttcacatcttccaataaaggtctgctgtggtgcgttgcgaatttatgagatcgtt	1037
Db	1020	ggggttggttccaatccaacagccaaggtccatgctcaatggtgtagaaggtgttaaaaaactt	1079
QY	1038	caatgcttccctactgctcttggtgtgtgtgtgttactactatccgcgaatggtgccggtt	1097
Db	1080	taacttaccatctactgctcttggaagggtggtgctacacaatccgtaaatggttcgcgagt	1139
QY	1098	ctggtgctaagagacggtgagttggaacttggagttggaagttggaagacaagatgcggagca	1157
Db	1140	ttgscatatatgagcgtcagttggtgccttgatttgtagatttcccaatgagttgcatataa	1199
QY	1158	tgaataatatagaataacttgggtccagaactataacttcaactcagttgctccaaagtaatgta	1217
Db	1200	tgatattactttagatatttggaacacagcttcaaacatgcatatattgcttccaataagac	1259
QY	1218	aaataagaattctcgctcagtgctctgaagagatctgcaatgaccttcccaaatctctc	1277
Db	1260	aaaccagaacatcccgagaatatatgtaaaagaataaaacacggttltglttaaaatttgcg	1319
QY	1278	taagcttcaagatcgtccaaagtttatacatttcaggagaagaacacgtgatacagaagctc	1337
Db	1320	catgcttacctatgcaactcgtgtgccaagatgcagaagctatttccaagaagaatgctgttcata	1379
QY	1338	cgaggttgatgagaagcaagaagatg	1363

Db	1080	taacttaccattactgattgcttggaggaggttgcctacaccaatccgtaatglttgcctgaty	1139
Oy	1098	cttgatgtctacgagacgtgagttgtcgaatttggagattgaagttaagacagaatgctgcggagca	1157
Db	1140	ttggacacatatagagctgcgaattggtcccttggatttggagattcccaatgaatttgcataaa	1199
Oy	1158	tgaatatatgaataactttgtcccaagctatcaacttcaagttgtctccaagaataatgta	1217
Db	1200	tgtattacttggagatttttggaccagagcttcaactcagatattatgcttctaactgaatgac	1259
Oy	1218	aaataagaattcttcgttcagatgcttgaagaagattcgaatgaacctcttcacaaactctc	1277
Db	1260	aaaccagaacacctccagatatataatggaanaagataaaacgcgcttgttgaaaatttgcg	1319
Oy	1278	taagcttcgaagctgtcccaagttgacatttccagaaagaaccgcctgaatacagaatccc	1337
Db	1320	catgttaccatcgtcacctcgtgtgcacagatgacagttattccagaagaatgctgttcatag	1379
Oy	1338	cgaggttgatgaagacaagaagaatg	1363
Db	1380	agacagttgagatgaagaatgtagaag	1405

RESULT 13

ID	AAC89555 standard; DNA; 1985 BP.
XX	
AC	AAC89555;
XX	
DT	08-MAR-2001 (first entry)
XX	
DE	Human histone deacetylase HDAC-2 coding sequence.
XX	
KW	Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KM	HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;
KM	gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200071703-A2.
PD	30-NOV-2000.
XX	
PF	03-MAY-2000; 2000MO-IB01252.
XX	
PR	03-MAY-1999; 99US-0132287.
XX	
PA	(METH-) METHYLGENE INC.
XX	
PI	MacLeod AR, Li Z, Besterman JM;
DR	WPI; 2001-016407/02.
DR	P-PSDB; AAB49955.
XX	
PT	Antisense oligonucleotide that inhibits expression of a histone
PT	deacetylase, useful for treating and/or alleviating the symptoms of
PT	neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX	
PS	Disclosure; Page 53; 125pp; English.
XX	
CC	The present invention provides inhibitors of histone deacetylase enzymes
CC	such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
CC	inhibitors may be antisense strands or they may be compounds identified
CC	by contacting the enzyme with the compound and measuring the resulting
CC	enzyme activity. These inhibitors are useful for treating cancers and for
CC	identifying which histone deacetylase is involved in a neoplasia.
XX	
SQ	Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

OY	138	ggatgatgaatacaggcgagcaatcgcgttcgctcgccgagcactgtgtgtgtgaagaggaagat	197
Db	180	gggtggcgcggtggcgggagggcccatggcggtacatgacatgaagagggcgcaadaaaaaaagt	239
OY	198	ttgtattctctatgacccttgaggttcgagcaattactactatgtgcgaaggtcatcccatgaa	257
Db	240	ctgtctactactcaagggtgatattgttgaaattattattattattgtgaacaggtcatcccatgaa	299
OY	258	gcccacatcgacatcccgatgaagcccaatgcctctcgtctcaactgaggtctcttcacagatat	317
Db	300	gcccataagaaatcccgatgaagcccaataactgtcgtttaaatatattgtcttatacagaanaat	359
OY	318	gcaagttctcaagacctccctcccgcggaagctgtactctgcccgttccacgcgcgcgcga	377
Db	360	ggaatatataatagcccacataaagcaactgcgcgaagaagaatgcaaaaatcatcagctgtaga	419
OY	378	ctatgtcctttctccgcagcatlaacccctgaagcccaagatcagatttcgccaact	437
Db	420	gtatatcaaatcttatcgcgtacataaagacagaataacatgctctgtagtatagtaaagcagat	479
OY	438	taagcgctcaatgttggtagaacgtgcgcgctctttagcggcctttatcttccctttgcc	497
Db	480	gcataatttaacgttggaaagaatgtgccagcggtttagaacctttgagtttgcga	539
OY	498	gacctatgcttggaggaactctgttggctctgttgaacgtcttaacacagcgcccttcgatat	557
Db	540	gctctcaacgtcgcggttcagttcggtgtgagtgagtaagttaacgcgacaaacagactgatat	599
OY	558	tgcctatcacacgggctgtgtgtgtctccatcagctaaagaatgtagcgagccttgctcttg	617
Db	600	ggctgttaattggctgtggaggaattacaactcgttaagaatacagaagcatcagagctctg	659
OY	618	ttacgctcatatctcgtctactagctacgcacgtcccttaacgacatgacatgactgtct	677
Db	660	ttacgcttaatgatattgtgtgtcgtaccccttgagattaccaaagtatcatctcagagagctt	719
OY	678	ttatgtcgaatatgataaccaccaagggaatgtgagtggaagacatttatgtactatga	737
Db	720	atatatgtatataagatatcatcatcatgatgtgtgttgaagaagcttttatacaacaga	779
OY	738	cagggttatgcctgtctcgttttataaatttgtttgtatcttcccggtlacaggtcatat	797
Db	780	tcggtgaatgcggtcatctccataataatgvgggaaatccttcttcgacacagagagact	839
OY	798	tcaggtataagtatgtatgtagcggaaagtactattctctcaatgttaccacatgtagatag	857
Db	840	gagggatattgtgtctggaagaagcaaatcattcatgttccaatlttccaatgtgtgtagt	899
OY	858	aatcgatgataagggcctatatctctgttatccaagcccacataggggaaagtlatggaat	917
Db	900	tataatgataagatcatatagggacatatttlaagcctatattccataaagggtatgagaat	959
OY	918	ttccgagcaaggcgctgtgttatgttcaatgttggatgttgatccctatcggggatcggtt	977
Db	960	gtatcaacactagtgctgtgtgtatttaacagtggtgtgtagcagctcatattatcgttgatag	1019
OY	978	aggttgtctcaactcttccaatcaaaaggtcatcgtgtggtgcgtcaaatltatagagatcgt	1037
Db	1020	gggtgttttccatctcaacacatcgaaggtctgcgttaaatgtttgtagaaggttttaaaaactt	1079
OY	1038	caatgttcccctactgtctctgtgtgtgtgtgttaccatctccgaatgtgcccgtt	1097
Db	1080	taacttaccatctactgtatgctgtggaggaagtggtctacacaacgcgttaatgttgtctcgat	1139
OY	1098	ctgtgttactcaagactcgtgaatgtgcaatttggagtgttgaagtltgaagaacaagatcgcgga	1157
Db	1140	tttgacataatgagcgtcaggtgcgccttgattgttgagatttcccaatgagtttgcacataaa	1199
OY	1158	tgaatatataatgaacttgtgtccagacataaacttcaactcagttgtctccaagtataatga	1217
Db	1200	tgattacttttgagttattgttgacacagcttccaacgtcatatatttccctccaacatgac	1259

OY	1218	aaataaataatctcgcagatcgcttgaagaagttgcgaatcaccttccacaaatctc	1277
Db	1260	aaaccagaacacccccaagaataataatgaaaagataaacacgcgtttgttgaanaattgcg	1319
OY	1278	taagcttcacgaatgcgtcccaagctgtacacatttcagaagaacacacctgatatacagaagctcc	1337
Db	1320	catgttaccctcatgacacccggtgtcagatgttgaagctattccagaagaatgtgtctatga	1379
OY	1338	cgaagttgatlgaagaccacaagaagatg	1363
Db	1380	agacagtgtagatgaagatgagagaag	1405
RESULT 14			
AAT86371			
ID	AAT86371	standard; cDNA; 1449 BP.	
XX			
AC	AAT86371:		
DT	16-APR-1998	(first entry)	
XX			
DE	cDNA encoding a histone deacetylase (HDx) protein, designated HD1.		
XX			
KW	Histone deacetylase gene; HDx; HD1; HDx polypeptide; deacetylation; H3;		
KW	H4; cell differentiation; chromatin structure; cell cycle progression;		
KW	proliferative disorder; fibroproliferative disorder;		
KW	degenerative disorder; autoimmune disease; HDx inhibitor; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..0	
FT		/*tag- a	
XX			
PN	MO9735990-A2.		
XX			
PD	02-OCT-1997.		
XX			
PF	26-MAR-1997;	97MO-US05275.	
XX			
PR	26-MAR-1996;	96US-0624735.	
XX			
PA	(HARD) HARVARD COLLEGE.		
XX			
PI	Hassig CA, Jamison TF, Schreiber SL, Taunton J;		
DR	WPI: 1997-489651/45.		
DR	P-PSDB; AAM29324.		
XX			
PT	New isolated histone deacetylase polypeptide(s) and genes - used to		
PT	develop products for modulating the proliferation, survival or		
PT	differentiation of cells, e.g. for treating tumours.		
XX			
PS	Claim 27: Pages 112-114; 159pp: English.		
XX			
CC	the present cDNA sequence encodes a novel histone deacetylase (HDx)		
CC	polypeptide, designated HD1. The HDx polypeptides are capable of		
CC	modulating proliferation survival and differentiation of cells. The		
CC	proteins are able to alter chromatin structure by deacetylating histones		
CC	such as H3 or H4. They have the ability to modulate cell growth by		
CC	influencing cell cycle progression or to modulate gene transcription. The		
CC	products can be used for diagnosis and therapy. They can be used, for		
CC	example, to treat tumours or proliferative disorders or spermatogenesis,		
CC	osteogenesis, chondrogenesis or the differentiation of progenitor cells.		
CC	They can also be used to treat psoriasis, bone diseases,		
CC	fibroproliferative disorders, degenerative disorders, or for repair of		
CC	cartilage, increasing bone density, liver repair subsequent to a partial		
CC	hepatectomy, to promote regeneration of lung tissue in the treatment of		
CC	emphysema, or for inducing tolerance in autoimmune diseases, and		
CC	transplant recipients. HDx inhibitors can be used as anti-fungal agents,		
CC	preservatives in foodstuff, feed supplements for promoting weight gain in		
CC	livestock, disinfectants, insecticides or defoliants. The products can		
CC	also be used in cell cultures.		

XX	Sequence	1449 BP; 406 A; 329 C; 396 G; 318 T; 0 other:
XX	50	
	Query Match	27.7%; Score 499.8; DB 18; Length 1449;
	Best Local Similarity	64.1%; Pred. No. 14e-142;
	Matches 753; Conservative	0; Mismatches 422; Indels 0; Gaps
QY	189	gaagaaagtttgatcttattcctatgacctgaagtcgcgaattactactatgacgaagta 248
DB	24	gaggaaagctctgttactactacgaacgggagtgttgaaattactatattgataaagca 83
QY	249	tcccatgaagcccatcgcatacgcgatgaaccatgcctctcctgcctcaatacgtctct 308
DB	84	cccatgaagcctccacgcgaatccgatcagatcataattgtctgcctcaatacgtctcta 143
QY	309	tcaagatagtcaggtctcgaagcctctccctccgcgcgaacgtgatactctgcgccttcca 368
DB	144	ccgaataaattggaatctatactgcacctcaaaaagcaatgtctgagagatgacccaatgaca 203
QY	369	cgcgcgaacatgatctctcttctccgcgaattaccctgtgaaccccgacaagatcagat 428
DB	204	cagagatgactactaattaatcttctgctccatcgcctccagatacaatgctcgagtaag 263
QY	429	tccgcaacttaagcgccttcaatglttggtgaagactgtccgcctcttgacagccttattc 488
DB	264	caagcagatgcagagatgaatacaagttgttgagagcgtccagatattcgatgcgtgttga 323
QY	489	ctttgcgaacctaatagtctggaagatcgtgttggtgctctgcaagcttaaccacgcct 548
DB	324	gtctctgcagctgtgtctactctgtgctctgttggaaggtgcttggaacttaataagcaaga 383
QY	549	ctgcgaatttgcacataacggcgctgtgtgtccctacacgcgttaagaagttggaagctc 608
DB	384	gacggacaatcgccttggaatcgtggcggtggggcgtgcacatagaaagaagtcggagctc 443
QY	609	tgcctctcgttaacgtcaatgatacgtctctaagctatacctaagcctcttaagcaagata 668
DB	444	tgctctcgttaacgtcaatgatacgtctcttgccatccgtgaactgttaagatatacca 503
QY	669	gcggtcttattatgtcatatttatatccacacgggagatgagtggaaggaacttta 728
DB	504	gaagggtctgtaacttgacattatataatccacatggtagcgcgtggaagaagccttcta 563
QY	729	tgcctcgaacgggttatgtactgtccggttataaatttggttgattacttcccggtac 788
DB	564	caccaacggaacgggtcatgtactgtgtcccttcttaagataagtagaagatctatcccgagac 623
QY	789	agttcacattcaagataataagttatgttagtcgaagaagtaacttctctcaatgtaaccat 848
DB	624	tgggaccttaagatatacgtggcgctgacaaagcaagatatattgctgtaactaccgt. 683
QY	849	ggaatgatggaatcgaatgaatgaagagctcaatctgttatcttaagcccatcatgaggaaagt 908
DB	684	ccgagacgggatttgatgaacgagctatagggccatttcaagccggtatgltccaaagt 743
QY	909	tatggaatttccgaccacggcgctgtgtgattgcaagtgtgtgtgcatactccatactcgg 968
DB	744	aattggaatggttcccaagccttagtcggtgtgtcttaacgattgtgttcgaagatccctatactcgg 803
QY	969	ggaatcggttagttgtcttcaactcttcaatcaagaagtaactgtctgagtggttaaatatt 1028
DB	804	ggaatcggttagttgtcttcaactcttaactatacaagaagacgcgcaagttgtgtgaatttgt 863
QY	1029	gagatcgftcaatgtctccctactgctcttggttggttggtgttacaatctccgaactgt 1088
DB	864	caagaagctttaaactgtcctactgctgattgctggaagcggtgtgttcaacattctaaagct 923
QY	1089	tgcgcgttgctgtggtcgaacgagagctgagttgactcttgagttgaagttgaagacaagaat 1148
DB	924	tgcgcgttgctgtggaacataatgaagagcgtgtgtgcccctggatacggagatctccataagact 983
QY	1149	gcgcgagatataatataataatatttctgttccagagctataacttcaacttcaactgttccaa 1208

Search completed: April 28, 2002, 19:28:40
Job time: 3564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 16:52:30 ; Search time 1866.22 seconds
(without alignments)
20262.483 Million cell updates/sec

Title: US-09-645-337-1

Perfect score: 1807
Sequence: 1 agagagcagctccctccccc.....atgatgatgatgatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenBml: 1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sy: *
12: gb_un: *
13: gb_vl: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgc_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1807	100.0	1807	6	AX139014	Sequence
2	1807	100.0	1807	8	AF195547	Arabidops
3	1807	100.0	1808	8	AF014824	AF014824 Arabidops
4	959	53.1	147215	8	AF120010	AL035538 Arabidops
5	959	53.1	199789	8	ATC81489	AL161593 Arabidops
6	795	44.0	1839	8	AF282858	AF282858 Mesembrya
7	740.4	41.0	1906	8	AF384032	AF384032 Zea mays
8	740.4	41.0	1943	6	AF168366	AF168366 Zea mays
9	714.6	39.5	1579	8	AF332875	AF332875 Oryza sat
10	709.6	39.3	2047	8	AF035815	AF035815 Zea mays
11	698.6	38.7	2019	6	AF168365	AF168365 Sequence
12	559.4	31.0	1742	8	AF045473	AF045473 Zea mays
13	559.4	31.0	1826	6	AF168363	AF168363 Sequence
14	557	30.8	1662	8	AF440226	AF440226 Zea mays
15	525	29.1	1704	8	AF072201	AF072201 Arabidops
16	523.4	29.0	1733	8	AF195548	AF195548 Arabidops
17	523.4	29.0	1800	6	AX139016	AX139016 Sequence
18	508.8	28.2	1997	6	AX305742	AX305742 Sequence
19	508.8	28.2	1997	10	MMU31758	U31758 Mus musculu
20	504.4	27.9	1985	6	AX053103	AX053103 Sequence
21	504.4	27.9	1985	9	HSU31814	U31814 Human traus
22	502.4	27.8	154409	2	AP003724	AP003724 Oryza sat
23	500.2	27.7	1977	10	MMHIDE	X98207 M.musculus
24	499.8	27.7	1611	6	AX053101	AX053101 Sequence
25	499.8	27.7	1611	9	HSU50079	U50079 Human histo
26	499.8	27.7	2093	9	BC000301	BC000301 Homo sapi
27	499.4	27.6	1999	10	MMO80780	U80780 Mus musculu
28	498.2	27.6	2091	9	D50405	U50405 Human mRNa
29	498.2	27.6	2111	6	AR012007	AR012007 Sequence
30	498.2	27.6	2111	6	162388	E11455 cDNA encodi
31	498.2	27.6	2111	23	E11455	162388 Sequence 2
32	495.8	27.4	2157	3	AY058487	AY058487 Drosophi
33	492.2	27.2	1646	5	AF043328	AF043328 Gallus ga
34	492.2	27.2	1646	5	AF044169	AF044169 Gallus ga
35	492.2	27.2	1678	5	AF039751	AF039751 Gallus ga
36	492.2	27.2	2305	5	XLAB21	X78454 X.laevus AB
37	489.4	27.1	1487	5	AF020658	AF020658 xenopus 1
38	483.2	26.7	2180	3	AF032919	AF032919 Strongylo
39	479.8	26.6	2172	3	DMH158AC	Y09258 D.melanogas
40	479.6	26.5	2073	3	AF026949	AF026949 Drosophi
41	474.4	26.3	1878	5	AF039752	AF039752 Gallus ga
42	473.6	26.2	1694	8	AF440228	AF440228 Zea mays
43	468	25.9	1576	6	AR168367	AR168367 Sequence
44	452.4	25.0	1799	10	AF321131	AF321131 Rattus no
45	451.8	25.0	1954	6	AX053105	AX053105 Sequence

ALIGNMENTS

RESULT	1	1807 bp	DNA	linear	PAT 30-MAY-2001
AX139014	AX139014	Sequence 1 from Patent EP1094112.			
LOCUS	AX139014	1807 bp	DNA	linear	PAT 30-MAY-2001
DEFINITION	AX139014				
ACCESSION	AX139014				
VERSION	AX139014.1	GI:14274698			
KEYWORDS					
SOURCE					
ORGANISM	Arabidopsis thaliana				
REFERENCE					
AUTHORS	Wu, K., Miki, B. L., Tian, L. and Brown, D. C.				
TITLE	Repressing gene expression in plants				
JOURNAL	Patent: EP 1094112-A 1 25-APR-2001;				
FEATURES					
source	1. 1807				
BASE COUNT	472 a	386 c	464 g	485 t	
ORIGIN					

TITLE Functional analysis of RPD3 histone deacetylase homologs in Arabidopsis thaliana
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1807)
 AUTHORS Wu, K., Tian, L., Malik, K., Brown, D. and Miki, B.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
 FEATURES
 Source
 1. 1807
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /clone="G11C3"
 gene
 1. 1807
 /gene="RPD3A"
 142. 1647
 /gene="RPD3A"
 /codon_start=1
 /product="putative histone deacetylase"
 /protein_id="AAC28474.1"
 /db_xref="GI:1106139"
 /translation="MDTGSNLASGPDVYKRYCYEPEVGNVYVGGHPMKPHIR
 MTHALLAHYGLQHMVLYKFPFARERDI CFHADYVSLRSTPTQDQIRQLRF
 NGEDEPFDGLQHMVLYKFPFARERDI CFHADYVSLRSTPTQDQIRQLRF
 VNDIVALELLKQHERVLYVDIDIHGGDGEAFATPDVMTVSHKXDPYRGCH
 IODIGSGKYYSLNPLDGDIDDESILHLFRIMKVMETIPPCAVLQCGDSLGS
 DRIGCTNLSTKGAECYKTRKSNVPLLLGGGGYTRNARWCETGVALSEVED
 KMPHEHYEYEGDYTLHVAIPSNMKNNSOMLEIRNDLHMLSKLOHAPVPEER
 PPDTEPDEVEDDEDKRPMDPSMDVDVDDRRIPSRVREAVPEPTDKDKLKGIM
 ERKGEVEVEDSGSTKVTGVPVVEEASVKKMEEGTNGKGEQAFPPKT"
 BASE COUNT 472 a 386 c 464 g 485 t
 ORIGIN
 Query Match 100.0%; Score 1807; DB 8; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 CTTATTCTTTTCCAGACCTATGCTGAGAGATCTGTGGCTCTGTCAAGCTTAA 540
 Qy 541 caagcctctgagatattgccatcaactgagctggtgtctccatcaagctaaagatgc 600
 Db 541 CACGGCTCTGCATATGTCATCAATCAATCGGGTGGTGTCTCCATCAAGCTAAAGTGC 600
 Qy 601 gaggcctctgctctgttaccatgagatgagatgctgttaccatgagatgagatgag 660
 Db 601 GAGCCTCTGCTCTGTGTTACGTAAGATATGCTTACTATCTTACGCTCTCTTAA 660
 Qy 661 cagcatgagatgttcttcttaccatgagatgagatgagatgagatgagatgagatgag 720
 Db 661 CAGATGAGGCTGTCTTATGTCATATGATGATGATGATGATGATGATGATGATGATGAT 720
 Qy 721 gcatltaagctactgacagagatgagatgagatgagatgagatgagatgagatgagat 780
 Db 721 GCATTTATCTACTGACAGGTTATGATGATGATGATGATGATGATGATGATGATGAT 780
 Qy 781 cccggtacagatgacatgacatgagatgagatgagatgagatgagatgagatgagat 840
 Db 781 CCGGATACAGGTACATTCAGATATAGCTTATGATGATGATGATGATGATGATGATGAT 840
 Qy 841 glaccatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 900
 Db 841 GTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 Qy 901 gggaaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 960
 Db 901 GGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Qy 961 ctatctgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1020
 Db 961 CTATCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Qy 1021 aaattatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1080
 Db 1021 AAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 Qy 1081 cgaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1140
 Db 1081 CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Qy 1141 gacaaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1200
 Db 1141 GACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Qy 1201 gttccaaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1260
 Db 1201 GCTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Qy 1261 ctctccaaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1320
 Db 1261 CTCTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Qy 1321 cctgataagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1380
 Db 1321 CCTGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Qy 1381 cccgattcagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1440
 Db 1381 CCGGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Qy 1441 gaaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1500
 Db 1441 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Qy 1501 aaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1560
 Db 1501 AAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Qy 1561 gtggagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1620
 Db 1561 GTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620

QY 1621 gagcaggccgttccctccctaaacaataagcctcgagcgttcaattctctgtacttttc 1680
|||||
Db 1621 GAGCAGGCCGTTTCTCTCTCTAAACATATAGACTCGGACCTTTTAATTTCTGCTACTTTTC 1680
QY 1681 tgctatcaaatgttgcataagtttcttgaggttctgttctgttgaagcactccctcg 1740
|||||
Db 1681 TGCTATCAAAATGTGTCTACTTAAAGTTTTCGAGTTGTTGTTGTGTGATGACATCCCTCTG 1740
QY 1741 ttttaaggagatbgaagcaggaatagtattatcttgltgcagtctgaaatgagataagat 1800
|||||
Db 1741 TTTTAAAGGAGATTGAGCAGCGATATGATTTATTGCTGATGCTCGAATGATGATATGAT 1800
QY 1801 atgacaa 1807
|||||
Db 1801 ATGACAA 1807

RESULT 3
AF014824

QY	61	agagatcatgcgaactctctctccgcgcacattgacgagacgtgattacaacacgct	120
Db	62	AGAGATCATGCGACACTTCTCTCCGACACATTTGACATGGGACACTGTGATTACAAACACCGT	121
QY	121	tgatcctacgaaaaagaagtaatgatalactgbcgagcaattcgctgagctccgacctgaa	180
Db	122	TGATCTTACGAAAAAGAGTAATGGAATGAGTACGGGCGCAATTCGCTGGCTCGGACCTGAT	181
QY	181	ggtgtgaagagggaagtgtgtattctctatgacctgaggtgtggaattactactatggc	240
Db	182	GGTGTGAAGAGGAAAGTTTGTATTTCCTATGACCCCTGAGAGGTGCGCAATTACTACTATGCC	241
QY	241	caaggtcatcccatgaaagcccacatgcacatccgatbaaccaatgacctctgcacacac	300
Db	242	CAAGGTCTATCCCATGAAAGCCCCATTCGCATTCGGATGACCCATGCCCTCTCGCTACTAC	301
QY	301	ggtctccttcagcatatgcaggtctcctaaagccctccctgcgcgagaaagtatctctgc	360
Db	302	GGTCTCTTACGATATGAGAGTGTCTCAAGCCCTTCCCTGCGCGGAAACGTATCTGTGC	361
QY	361	cgcttccagcgcgagacatgtctcttctccgcgacgataccccgaaacccagca	420
Db	362	CGCTTCCAGCGCGACGACATGTCTCTTTCTCCGACGATTACCCCTGAAAACCCAGCA	421
QY	421	gataagattccgcaactlaagccgttcaaatglttggtagaacgtgcctgtctttagagc	480
Db	422	GATAGATTCCCAACTTAAAGCGCTTCAATGTTGGTGAAGACTGTCCGCTCTTTACGCGC	481
QY	481	cttattcctcttgcagaccctatgctctgagagatcgttctgtgctctgtcaacttaac	540
Db	482	CTTATTCTCTTTGCCGACCTATGTGTGGAGATGCTGTTGGTGGCTGTGCAAGCTTAAC	541
QY	541	caagcctctgagatatctgccatcaactgagctggtgtgtctccatcaagctaaagaagtgc	600
Db	542	CAGCGCCTCTCGCATTTATGCCATCAACTGGCGCTGGCTCTCATACAGCTAAGCAAGTGC	601
QY	601	gaagcctctgagctctgttaccgtcaatgatatgcgtcttagctatctcctagacgtcctlaa	660
Db	602	GAGGCTCTGTGGCTTCTTTACGTCAATGATATGCTTTAGCTACTTCTAGAGCTCTTTAAG	661
QY	661	cagatgagcgtgtcttctttagtcgatalatgatalcaccaacagggagatgagatgagag	720
Db	662	CAGATGAGCTGTGTTTATGTCGATATTGATATCCACACGGGAGATGAGATGGAGAG	721
QY	721	gcaattatgctactgcagaggttatagactgctcgttccataaattgtgtatcttct	780
Db	722	GCATTTTATGCTACTGACAGGGTTATGACGTCTGCTTTCAATAATTGGTGATATTCTTT	781
QY	781	cccgataagctacatctcagagataagttatagttagtagaggaatfactactctcaat	840
Db	782	CCCGTACAGGCTCACATTACGATATGATAGGTTATGATAGCGGAAGTACTATTCTCTCAAT	841
QY	841	gtaccactgagatgtagaatacgatagatgagagctatcatctgttatttlaagcccacatg	900
Db	842	GTAACACTGAGTATGATGAATTCGATGATGAGAGTATCATCTGTATTTCAGGCCATTATG	901
QY	901	gggaaagtatggaattctccgaccaggggctgtgtatttgcnaatgtgtgtgactgc	960
Db	902	GGGAAAGTTATGAGAAATTTTCCGACCAAGGGGCGTGTGTTATTCGAATGTGCTCATGCC	961
QY	961	ctatctgggagatcggttaggtgtgtcttcaaatcttcaatcaaaaggtcagtcgagtggtc	1020
Db	962	CTATCTGGGAGTCGGTTAGGTATGCTTCAACTTTTCATATAAAGGTCATGCTGAGGCGTC	1021
QY	1021	aaattatgatatggttcaatgttccctactgctcttggtgtgtgtgtgttlaacatac	1080
Db	1022	AAATTTATGAAATGCTTCAATGTTCCCTTACTGCTCTTGGGTGGTGGTGTATACACTATC	1081
QY	1081	cgcgaatgttcgcgctgtctgtgtgctacagaaacgagatgtgcaccttggagttaagttaa	1140
Db	1082	CGCAATGTGTCCGCTGTGCTGTGTGTACGAAAGCTGGAGTTGCACCTTGGAGTTTAACTTGA	1141
QY	1141	gacaaagtcgcgagcatataatlatgaalacttgtgtccagacataacactcagct	1200

OLMSSKSKSLGFTSLSFSEFLSPYTKLSVEKLELTPTKRGKADDPNPLSLRAVAST
 PPTRSPTVHLVQARSLKSGSELPMLRVRRLAKSPTRVADDETETVNLQMRNSDA
 AFGDRQVYIVGKGEETVLAEDGTSLSDSKWSLKOTCNPATDGLFELSGTRM
 VVYSGRKLSEPKHCKSLRSBODPMTAVEFSKQHPYGAVALDLKFGSIEANKML
 VLPQWSSFLILSDLLKKEGSAAKDQIVANGITEESREIVLISOEKLSEEMMVD
 TTPVAVAARKINGARCSKELSGNNIEPEGHCGCCGCCGCCGCCGCCGCCGCC
 TKIEGGGGSGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 complement(10265..10337)
 /gene="F20D10.20"
 /number=1
 complement(10338..11657)
 /gene="F20D10.20"
 /number=2
 complement(11658..11740)
 /gene="F20D10.20"
 /number=2
 complement(11741..11989)
 /gene="F20D10.20"
 /number=3
 complement(11990..12070)
 /gene="F20D10.20"
 /number=3
 complement(12071..12191)
 /gene="F20D10.20"
 /number=4
 complement(12192..12299)
 /gene="F20D10.20"
 /number=4
 complement(12300..12463)
 /gene="F20D10.20"
 /number=5
 complement(12464..14030,14118..14738,14828..15394,
 15491..15569,15669..15730,16043..16069)
 /gene="F20D10.30"
 13386..16069
 /gene="F20D10.30"
 complement(16070..16386,16403,16418..16738,16828..17394,
 17491..17569,17669..17730,18043..18069)
 /gene="F20D10.30"
 /note="strong similarity to dnaK-type molecular chaperone
 HSP1-plum sativum, PID:q20835
 contains ATP/GTP-binding site motif A (P-loop) (AEAVLGKS),
 heat shock hsp70 proteins family signatures (IDLGTSN),
 hsp70_2 (VFDLGGTDFVSL), hsp70_3 (VLLVGGMTVPKVE),
 contains EST gb:U4452, Z24523, T04782, Z17893, Z17973"
 /codon_start=1
 /product="heat shock protein 70 like protein"
 /protein_id="CA33753.1"
 /db_xref="GI:4467097"
 /translation="MASVASFKSVSANGNSMFKGLYLARPCSRPQNDVIGDLG
 TNSCVSMGKTAARVTEAEGSRTPSVVANNOKGLVGPARKQAVTNPNTIG
 SKRLGRDDPQOTQKEMKVPYIKAPNGDAWANGKFSPOQAVNTKMKET
 AEAYLGKSINKAVTAVPAYFNDARQATDAGKLAGVQRIINPTAALVYGMNK
 EGVAVFDLGGTDFVSLIEISGVEVATNGDPLGGEDDNTLLELVNFRSD
 NIDLTPDNALQRLRAAKAKIELISTTQTEINLPITADSGAKHLNITVTRKE
 GLVGLITERSPQMLNDAGVTIKEDVEVLVGQMTVPVQELVSEIFSKSPCKG
 VNPDEAVAMGAALGGILRGDVKDLLDLVPLSLGIELGVAFKLIPRNTIPKPK
 SOVFTSAADNOMQVIGKVLQGRKEMADAKVIGEDPLVLPAPRQMPQIEVTFIDA
 NGITVSADKAKTEONITIRSSGSLSDIEINRWKVELNAOKQOEKQIDLRNS
 ADTITVSEKSLSEVEREKIPAEIASRLEFVADLRTAMGDEVEDIKAAVEANNAVS
 KIGEHMSKSG
 complement(13386..14030)
 /gene="F20D10.30"
 /number=1
 complement(14031..14117)
 /gene="F20D10.30"
 /number=1
 complement(14118..14738)
 /gene="F20D10.30"
 /number=2
 complement(14739..14827)
 /gene="F20D10.30"
 /number=2

exon complement(14828..15394)
 /gene="F20D10.30"
 /number=3
 intron complement(15395..15490)
 /gene="F20D10.30"
 /number=3
 exon complement(15491..15569)
 /gene="F20D10.30"
 /number=4
 intron complement(15570..15669)
 /gene="F20D10.30"
 /number=4
 exon complement(15670..15730)
 /gene="F20D10.30"
 /number=5
 intron complement(15731..16042)
 /gene="F20D10.30"
 /number=5
 exon complement(16043..16069)
 /gene="F20D10.30"
 /number=6
 gene 16556..19503
 /gene="F20D10.40"
 exon complement(16556..16790)
 /gene="F20D10.40"
 /number=1
 gene complement(16556..16790,16869..17056,17133..17377,
 17453..17618,17710..17975,18152..18358,18789..19503)
 /gene="F20D10.40"
 complement(16556..16790,16869..17056,17133..17377,
 17453..17618,17710..17975,18152..18358,18789..19503)
 Query Match 53.1%; Score 959; DB 8; Length 147215;
 Best Local Similarity 84.3%; Pred. No. 1.1e-258;
 Matches 1199; Conservative 0; Mismatches 10; Indels 214; Gaps 2;
 QY 135 agaggtgaatgagatctgagcgcaattcgctgctgcgcggaacctgagtgtgtaagaagaa 194
 DB 87082 ACAGGTAATGATGATACGCGGGAATTCGCTGCTCGGACCTGATGCTGAAGAGAA 87023
 QY 195 agttgtattctatgagccctgagcgtaattactactatgagcgaagtcacccat 254
 DB 87022 AGTTTGTATTCTATGACCTGAGCTGAGTGGCAATTCTATGCTATGCGCAAGCTATCCAT 86963
 QY 255 gaagcccaatgcacatgcagatgaccatgacctctgctgctacacgctcctcaaga 314
 DB 86962 GAAGCCCATGCGATCGCATGACCATGCGCTCTCGCTCACTAGCGTCTCTTCAGCA 86903
 QY 315 tatgcaagttcgaagccctcctcctgcccgaagtgatctctgccccttcaagccga 374
 DB 86902 TATGCAAGTTCTCAAGCCTTCCCTGCGGACCTGTATCTCTGCGCTTCACGCCGA 86843
 QY 375 caactatgctctctctcctgcagcatcaccctgaaccgaagaatcagatcagca 434
 DB 86842 CGACTATGCTCTCTTCTCTCGGACGATTACCCCTGAACCCGAAAGATCAATTCGCCA 86783
 QY 435 acttaagccttcaatgctggtgaagaagctccgctccttgaagcgccttattcctttg 494
 DB 86782 ACTTAAGCGCTCAATGTTGTTGAAGACATGTCGCCCTTGAAGCGCCTTATTCTTTTG 86723
 QY 495 ccaagcctatgctgagagatcgtgtgtgctgtgcaagcttaacacagagcctctgga 554
 DB 86722 CCAAGCCTATGCTGAGAGATCGTTGTGTGCTGTGCAAGCTTAACCAAGCCTCTGCGCA 86663
 QY 555 tatgccaatcaatgagcgtggtgtgtcctcaagcttaagaagtgagagcctctgact 614
 DB 86662 TATGCCATCACTGCGCTGCTGTCTCATACGCTTAAGAGTGGAGGCTCTTGCGCTT 86603
 QY 615 ctgttaagctcaatgacatgctcttaagcttaagccttcaagcgaagcgtg----- 667
 DB 86602 CTGTTAAGCTCAATGATGCTGTCTTACTATCTAGAGCTCTTAAGCAGCATGAGGTTTG 86543
 QY 668 ----- 667

/gene-"At4g38130"
 /number-4
 complement(6329, .6583)
 /gene-"At4g38130"
 /number-4
 complement(6584, .6814)
 /gene-"At4g38130"
 /number-5
 complement(6815, .6944)
 /gene-"At4g38130"
 /number-5
 complement(6845, .7385)
 /gene-"At4g38130"
 /number-6
 complement(7386, .7469)
 /gene-"At4g38130"
 /number-6
 complement(7470, .7997)
 /gene-"At4g38130"
 /number-7
 8808, .9245
 /gene-"At4g38140"
 complement(8808, .9245)
 /gene-"At4g38140"
 /number-1
 complement(8808, .9245)
 /gene-"At4g38140"
 complement(8808, .9245)
 /gene-"At4g38140"
 /note-"similarity to RING-H2 zinc finger protein RHA1a
 -Arabidopsis thaliana, PID:g3790554"
 /codon_start-1
 /product-"putative protein"
 /protein_id-"CA880479.1"
 /db_xref-"GI:7270798"
 /translation-"MISMLFRRSPICATAIVFYCYVICIPLGRKKKGADADDDG
 NLVGMGDKKEEETICCPICIVFEEDVATHLPKCAHLFINCIETPLMRGHLTGP
 LCRSFVLAPTPPTQNNNAHSSSTLTLSTFFFCIFLHLGTL"
 complement(10151, .11059)
 /gene-"At4g38150"
 10151, .11059
 /gene-"At4g38150"
 complement(10151, .11059)
 /gene-"At4g38150"
 /number-1
 complement(10151, .11059)
 /gene-"At4g38150"
 /note-"similarity to leaf protein (PNC401) -Pharbitis nil,
 PID:d1026846"
 /codon_start-1
 /product-"putative protein"
 /protein_id-"CA880480.1"
 /db_xref-"GI:7270799"
 /translation-"MYPSSKAVVFARQAKOIKVITTPSKATRFLLSTGDMGVDEQON
 PEPPLPKRPLGRSSSHREPPARQAHNLGKSDTLLSDGFLDEPKLVNDISRTLP
 KPPDLPDPPEDDELFPKMKEGGLIPNAAMLDDGLKDLVDEAKLFTLMDDK
 LGIPVAVIYAVVAFCAKIKEDAKRIFKMQNNQIAPAFSGVGLVGVGICVNCMLD
 DVAFCSEMLSGSHPNVPEVELVADLCVKVQDAQSIDTLNKKGFVAVNKAVKE
 PMDKRAPFPLAEAFIFKKKPTKPF"
 11352, .12353
 /gene-"At4g38160"
 11352, .12353
 /gene-"At4g38160"
 /number-1
 11352, .12353
 /gene-"At4g38160"
 /note-"contains EST qb:A1992903.1, H37273, A1100213,
 T41873, AA067478, A1100538"
 /codon_start-1
 /product-"putative protein"
 /protein_id-"CA880481.1"
 /db_xref-"GI:7270800"
 /translation-"MEVTNTSITMFFRDKGDDPSIDKMLRCKQLEKAQSDVASEN

MDYLSNIVGIORLDPYVSRCKPILTLRLDERLLPWECUSSLGRNPREVASATTKP
 PPILSHVEEKLCPILAFALGVPELOLGMILFNPRLISYIDTKLTVISPLASL
 GLDODMIGKLVKNNPFLMGYSVDRKRLRPTTEFLISVGLSDGIKSVAMPOLLCR
 DVNKLKLPYDYLKEGFGDSOIAITNVGYPQILIKVKNLSIOPLIRLVOMGMGM
 EVASYPEFFHNLKKKVESRFLVKVKNNDICSLRMLDCCNKKFKFKGFEVYASF"
 13703, .15379
 /gene-"At4g38170"
 13703, .15130
 /gene-"At4g38170"
 /number-1
 join(13703, .15130, 15212, .15379)
 /gene-"At4g38170"
 /note-"Contains prokaryotic membrane lipoprotein lipid
 attachment site AA73-83
 contains EST qb:N37149, T14014"
 /codon_start-1
 /product-"hypothetical protein"
 /protein_id-"CA880482.1"
 /db_xref-"GI:7270801"
 /translation-"MSRVHVLNLYLKRHLENPGLVLAIEDCGNFWFADPTCRILNT
 YFGDLVFDPTTYRRGRKRYOVPPAFTGFENHGOPLFCALILNESSESPALPOTWL
 QAMSAAPPESITVEPDRLOVAVSRVFSOTRFRSOPILFEETEXKLAVHQAHPTE
 SEFINCVETETRAAEFASMDSTVRYRTHMEDMDLOSTYNARQVRYFIRDTFTEL
 STNEGSLINSFQGVDASTVMQMLIOYERAIQDSMEKELKADYEATNSTPVKTP
 SPMEKQASLYTRAAFLKROEEFVELAIIPANIISDSOTHTTYRAKGEVAKHTVS
 PSLKVKANSCOMFEYSGITCRHILAVFASANNVLAISRYLIRMTKEAKIRGTEQ
 PFSNCGCESLNCFSNLEKATKVEEAKSIQLYKMAALDDEAKKVAASNRTP
 GTRLNGEAYPSEEARETANINHPGGEHERILBLTVELEBTGRCVRYANLILSL
 RMDEQRQLSTKVNARSLKE"
 15131, .15211
 /gene-"At4g38170"
 /number-1
 15212, .15379
 /gene-"At4g38170"
 /number-2
 complement(15642, .15815)
 /gene-"At4g38180"
 /number-1
 15642, .18344
 /gene-"At4g38180"
 complement(15642, .15815, 16152, .18344)
 /gene-"At4g38180"
 complement(join(15642, .15815, 16152, .18344))
 CDS

Query Match 53.1%; Score 959; DB 8; Length 199789;
 Best Local Similarity 84.3%; Pred. No. 1.2e-258;
 Matches 1199; Conservative 0; Mismatches 10; Indels 214; Gaps 2;

QY 135 aagagtaatgatalactgagcaatcgcgctgcgcgagcagatgagtgatcgtcgcgca 194
 |||||||
 Db 8004 AGAGTAATGATGATGAGCGGCAATTGCGTGGCGGACGACGATGATGATGAGCA 7945
 |||||||
 QY 195 agttgtattctatgacccctgaagtcgcaattactactatagccaagtcacccat 254
 |||||||
 Db 7944 AGTTGTATTCTATGACCCCTGAGGCTGCGCAATTACTATAGGCAAGGATATCCAT 7885
 |||||||
 QY 255 gaagcccaatcgatccgcaatgacccatgcgcctcgcgctcactaagatgctccatga 314
 |||||||
 Db 7884 GAAGCCCAATCGATCCGCAATGACCATGCGCTCGCTCGCTACTAGGTCCTCTTAGCA 7825
 |||||||
 QY 315 tatcaggtctcaagccctccctgcgcgagaaagtgatcgtcgcgtctccacgcga 374
 |||||||
 Db 7824 TATGACAGTTCTCAAGCGCTTCCTGCGCGGACGAGATGATGCTGCGCTTCACGCGCA 7765
 |||||||
 QY 375 cgaatagctctctctcgcgagcatcacccttgaaacccagcaagatcagatccga 434
 |||||||
 Db 7764 CGACTATGCTCTCTTCGCGAGCATATACCTGAAACCCACAGATTCAGATTCGCA 7705
 |||||||
 QY 435 acttaagcgttcaatggttgaaagcgtcgcgctcttgaagcgtctatccctttg 494
 |||||||
 Db 7704 ACTTAAAGCGTTCAATGTTGTGTAACGACTCTCCGCTTTGAGGCGCTTATTCCTTTG 7645
 |||||||
 QY 495 ccagacctatgctgagagatcgttggtgctctgtaacagatcaacacgacctcga 554

```

Db 7644 CCAAGACCTATGCTGAGAGATCTGTGGTGGCTGTGCAACCTTAACCAAGGCTCTGGCA 7585
Oy 555 tattgcataactatggcgtgtgtgtctccatcagctaaagaatggaagcccttgctt 614
Db 7584 TATGGCCATCAACGTGGCTGTGGTCCATCCAGCTTAACAAAGTGGAGGCCCTGTGGCT 7525
Oy 615 ctatcagctaatgatactgtcttaagctatccctagaagctccttaagaagatg----- 667
Db 7524 CTGTATAGTCAATGATATCTCTTACCTATCTAGAGCTCTTAAAGCAGCATAGGTTTG 7465
Oy 668 ----- 667
Db 7464 TCCACCTGACCGCTTGTGATATGCTAGAGATGTTATGCTGTTCTGAATCTTAAT 7405
Oy 668 ----- 710
Db 7404 CTTCCTCTCTCGTTTTCAGAGGTGTCTTTATGTCGATATGATATCCACCGGGGATGG 7345
Oy 711 agtggaggaagcaatltatgtctactgacaggtatgaagctgtctgtcttcataattgg 770
Db 7344 AGTGAAGAGAGCAATTTATGCTACTGACAGGTTATGACTGTCTGTTTCAATAATTTGG 7285
Oy 771 tgattacttcccgtaacaggtacacatlcaggatataagttatgtagcggaaagta 830
Db 7284 TGATTAATCTTCCGGTACAGGTACATTCAGATATAGTTATGTTAGCGGAAGTACTA 7225
Oy 831 ttctctcaatgtaccatgtagatgtagaatcgatgtagaagctatcatctgttatcaa 880
Db 7224 TTCTCTCAATGTCACATGTCATGATGATGATGATGATGATGATGATGATGATGAT 7165
Oy 891 gccatcatggaagaatltatggaatlttcgacacagagagctgtgtagttagcaatgtg 950
Db 7164 GCCATCATGAGGGAAGTATGGAATTTTCCACCAAGGAGGCTGTGTTATGGAAGTGG 7105
Oy 951 tgcctacccatctgtaggaatcggttaggtgtgtcttcaatccttcaatcaaaagctatg 1010
Db 7104 TGGTATATTCATGTCGTGATAGTGTGGGTGCTTTAATCTTCAATCAAGAGTCATGC 7045
Oy 1011 tgggtgtgtaaatatataagatcggttcaatgttccctacgtctctgtgtgtgtgtg 1070
Db 7044 TGAATGGGTAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 6985
Oy 1071 ttcaactatcgaagatgtaccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1108
Db 6984 TTTCACATATCCGCAATGTTGCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6925
Oy 1109 ----- 1108
Db 6924 TATGCTGCTAGCTTGTGATGAAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 6865
Oy 1109 ----- 1120
Db 6864 TGAAGCCTCTCTTTTGAATTTATTACAACAACCTGTTTATATGACCAACCTGGAGTTG 6805
Oy 1121 caattgagtgatgagtgaaagacagatgacgagagcaatgaatataatgaatcttggtc 1180
Db 6804 CACTTGAAGTGAAGTGAAGGACAAAGTCCCGAGCATGATGATGATGATGATGATGATG 6745
Oy 1181 cagactatacaacttcaactgtgtctcacaaglaacatggaataaagaatctcgtcgaatg 1240
Db 6744 CAGACTATACACTTACAGTGTGCTCCAGTAACATGGAATAATGAATTTCTGTCAGATGC 6685
Oy 1241 ttgaagagatctgcaatgacctctccacaatctctcctaagctcagcatgctccaagt 1300
Db 6684 TTGAAGAGATTCGCAATGACCTTCTCCACAACTCTCTTAAGCTTCAGCATGCTCCAAGTG 6625
Oy 1301 taccattcaggaagaagacacatgatatcagagactcccgagat 1343
Db 6624 TACCATTTCAGAAAGACCACTGATACAGAGACTCCCGAGGT 6582

```

RESULT 6

```

AF282858
LOCUS AF282858 1839 bp mRNA linear PLN 11-JUL-2000
DEFINITION Mesembryanthemum crystallinum histone deacetylase (Hdeac1) mRNA.
ACCESSION AF282858
VERSION AF282858.1 GI:9022438
KEYWORDS common ice plant.
SOURCE Mesembryanthemum crystallinum
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 1839)
AUTHORS Scharte,J. and Baur,B.
TITLE Molecular cloning of histone deacetylase from Mesembryanthemum
crystallinum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1839)
AUTHORS Scharte,J. and Baur,B.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) Institut fuer Botanik, Universitaet
Muenster, Schlossgarten 3, Muenster 48143, Germany
FEATURES
source
1..1839
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/lissue_type="leaf"
1..1839
/gene="Hdeac1"
174..1676
/gene="Hdeac1"
/codon_start=1
/product="histone deacetylase"
/protein_id="AAF82385.1"
/db_xref="GI:9022439"
/translaton="MDTGSNSLPSPGDCVKKRKYFYDPEVGNYYGCGHPMKPHRIR
MTHALHAYGLQHMVLPKVPARBRDLCRFHADYVSEFLNGIIPMEMOQDLRLKRP
NVGDCPPEDGLYSRCORYAGSGVAGAKLHIGCDIAIMNAGLHAKKKEAAGRCY
VNDIYLALBELKVERVLYDIDIDHSDCYBEAFTTERMTVSFHFQGYPFGETR
FVYIDGKATYSLNVPDLDIGDSDSGLKPLKMYELFKPGAVLQDGAOSLSD
DRUGFNLSLNGHACVYMSFNPVNLVGGGTYTJNVVARCMYETGVVALGPELD
KIPQHEVYEYEGPDYTLHVPASNMENKSRMLDIDKGLLEYLSLSQHAHSVQFOER
PPEAEIPEDEDOHDPEGRMPDPSDMEVDSDSISRSVRKEIREPDLKDDVDKDAY
DHPRMDVSYGAGNYSVAKPTTSSDQPEGCEMEDNSNRRPDGPNMNS"
BASE COUNT 468 a 383 c 452 g 536 t
ORIGIN
Query Match 44.0%; Score 795; DB 8; Length 1839;
Best Local Similarity 72.1%; Pred No. 7.8e-213;
Matches 1035; Conservative 0; Mismatches 400; Indels 0; Gaps 0;
Oy 91 ttgactgcagactgtatcaaacacacgcttgatcctacgaaagaaglaatgatact 150
Db 123 TTTTTCGGGCGACTTCAATTTTTCGGCTAGCCAGCCACTGTTTGGTATGATGATAC 182
Oy 151 ggcgcgaatctgcgtgcgtccggaactgagtgtgaaagaagaagttgtattctat 210
Db 183 GGAGCAATTCATCTTCATCTGACACGATGGGTAAACGTAACCTCTGACTTTTAT 242
Oy 211 gaccctgagtgcggaattactactatggcacaaggtcatcccatgaagcccatgcgcatc 270
Db 243 GATCCAGAACTTGGAACATTAATTAAGGCAAGGTACCTATGAAAGCCACATCGAATT 302
Oy 271 cgaatgaccatgccctccctcgtctcaactaagctcctcctcaacatagatgagttc 330
Db 303 CGTATGACAGATGCTTCTAGCTCACTATGGGCTCTTCAACATATGCAAGTCCTGAG 362
Oy 331 cctctccctgcgcgcgaacgtatctctgcgcttccacgcgcgaacatagtctcttt 390
Db 363 CCAGTCCCTGCAAGAGATAGAGACTTGTGAGGTTCCACGCTGATGATGATGATGATG 422
Oy 391 ctccgcagactaacccctgaanaccagcaagatcagattcgcaacttaagcgttcaat 450

```


QY 316 atgcaagttctcaagccctccctgcccgcgaacgtagctctgcgcgttcccaagccgac 375
 124 CTGAGAGTCTCCGCCCTACCGGCTCTGAGAGCGGACACTCCGCCGCTTCCACTCCGAC 183
 QY 376 gaactatgtctcttcccgcaagcattaccctgaacccgaacgaatcagattgcacaa 435
 184 GACTAGCTGCTTCTCCGCGGCTCCGACCGAACCCTGGGTCTCTCCGACCGCGCGCC 243
 QY 436 cttaagcgtctaatgttggtagaagctccgctctttagagcgtcttattcttcttgc 495
 244 ATTAAGGCGCTTAACGTCGCTGAGAGCTCCCGCTGTCCAGGCTCTTCCCTCTCTGC 303
 QY 496 cacacatctatgagatctgttggctgtcgtcaagcttaacacagcgccttgagat 555
 304 CAGGCTCCGCTGGGGGAACATCGGCGCGCGCTCAAGCTTAACCGCGGGAGCGCGAC 363
 QY 556 atgcacatcaactcgtggtgtgtctccatcaagcgttaagcgtcgaagcctctgtctc 615
 364 ATCAGCGTCAACTGGGCGGCGCGCTCCACACGCGCAAGAACGACGCGCTCCGGGTTG 423
 QY 616 tttatgctaatgatatgtcttagctatcttagaagctccttaagcagatgagtggt 675
 424 TGCTAGCTCAAGACATCGTCTCGCATCTCGAGCTCTCTCAAGTTCCACAGCGCTGTG 483
 QY 676 cttaagctcagatatgatatccacacagcgtgagtgagtgagcagcatltagtact 735
 484 CTATATGATGACATGATGTCCACACATGAGATGCGCTGAGAGAGCGCTTCTTCACTACA 543
 QY 736 gacagaggtatgacgtctcgtttcacaatttgtagattcttcccgtaacaggtcac 795
 544 AACCGAGTCATGATGCTTCTTCTTCAACATGAGGAGATTTTCCCTGGTACTGACAT 603
 QY 796 atccagatatagttatgtagtgaagaaagtactattcttcaatgtaccacgtgagat 855
 604 ATACTGACCTTGGGCGACCGGAGAACATTAATGCTGATGCTTCCCTGAGATGAT 663
 QY 856 ggaatcagatatgagcgtatcatctgttatacagcccatctcgtgggaagttatggaa 915
 664 GGTATGATGATGACACACACTTCTGTGAGGAGGTCATGCTGACTCCCTCCGTTCTTACGCTG 723
 QY 916 atttccgacagcagcgtctgtgtatgtgcaatgtgtgtgtgtgtgtgtgtgtgtgt 975
 724 GTTATTCAGGCAACAGCGGTTGCTCTCCATGCGGAGCTGACTCTTGGCTGGAGACAG 783
 QY 976 tttagttgtctcaatcttcaatcaaaagtcagtcgtagtggtcgaatatttagagatcg 1035
 784 TTAGGTTGCTTCAACCTGCTGTGAGGAGGTCATGCTGACTCCCTCCGTTCTTACGCTG 843
 QY 1036 ttcaatgtccctactgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1095
 844 TACAATGCTCTTGAATGATGCTTTAGTGTGTGAGGTTACACATGACAAATGTTGCAAGC 903
 QY 1096 tgcgtgtgtcagagagctgtgagttgcaactgtgagttgaaagtgaagacaagatgcgag 1155
 904 TGCCTGTGCTGACAGACAGCGGAGTGTGCTGTGTGAGATTGCAACCTGATTAACAAGCTTAC 963
 QY 1156 catgaatattgaatgaatcttggtagaagcttcaatcaatcaatcaatcaatcaatcaat 1215
 964 AATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
 QY 1216 gaaataagaattctcgtcagatgctgtgaagagattcgaatgaatcaatcttccacaatctc 1275
 1024 GAAATCCTGAATGACCAAAAGGACTTGGAGACATTAAGAACATGATATTTGGAAGAACTG 1083
 QY 1276 tctaaagctcagatgctcagatgtacacattcgaagaaagcaccgtatagagat 1335
 1084 TCAAAAGTAAAGACATGTTCCAGACACTCAATTCATGACAGACGCTGAGACCTGAAGCT 1143
 QY 1336 cccgaggttgatgaaga 1352
 Db 1144 CCAAGGAGAAAGAGA 1160

RESULT 15
 AY072201
 LOCUS
 DEFINITION
 Arabidopsis thaliana putative histone deacetylase (At5g63110) mRNA.
 complete cds.
 ACCESSION
 AY072201
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1704)
 2 (bases 1 to 1704)
 Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
 Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
 Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
 Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
 Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
 Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
 Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
 Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
 Full length cDNA of gene At5g63110 (GI:15242626)
 Unpublished
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT
 Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 RIKEN genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-length cDNA'; Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.)

The Sak, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
 Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
 Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.,
 Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
 Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
 Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
 Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
 this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
 contributed equally to this work as pls.

FEATURES

source

1. 1704
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL09-59-F23 (R12957)"
 /note="this clone is in a modified pBluescript vector
 (pUC-1) as a BamHI/XhoI insert.
 ecotype: Columbia"
 1. 1704
 /gene="At5g63110"
 1. 75
 /gene="At5g63110"
 76. 1491
 /gene="At5g63110"
 /codon_start=1
 /evidence=experimental
 /product="putative histone deacetylase"
 /protein_id="AAL60022.1"

CDS

gene

5'UTR

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:51:32 ; Search time 44.51 Seconds

(Without alignments)
1646.426 Million cell updates/sec

Title: US-09-645-337-2

Perfect score: 2722

Sequence: 1 MDGTGNSLASPGDVKRKVC.....KMEERGTHKGAEAFPEPT 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	99.9	501	10 Q9SLZ3	Q9SLZ3 arabidopsis
2	2134.5	78.4	500	10 Q9LKG1	Q9LKG1 mesembryant
3	1958.5	72.0	493	10 Q9AXFO	Q9AXFO oryza sativ
4	1568.5	57.6	471	10 Q9PML2	Q9PML2 arabidopsis
5	1561.5	57.4	471	10 Q9FVE5	Q9FVE5 arabidopsis
6	1548	56.9	458	10 Q9ZTP8	Q9ZTP8 zea mays (m
7	1496.5	55.0	521	5 Q9ZTP8	Q9ZTP8 zea mays (m
8	1496.5	55.0	521	5 Q9VZAL	Q9VZAL drosophila
9	1374	50.5	428	11 Q9PFA0	Q9PFA0 rattus norv
10	1361.5	50.0	687	3 Q9PFA5	Q9PFA5 emeticeila
11	1359	49.9	428	11 Q9PFA0	Q9PFA0 mus musculu
12	1358	49.1	405	3 Q9JLX5	Q9JLX5 mus musculu
13	1336	49.1	405	3 Q9JLX5	Q9JLX5 mus musculu
14	1334	49.1	405	3 Q9JLX5	Q9JLX5 mus musculu
15	1329.5	48.8	648	3 Q9CIC6	Q9CIC6 caenorhabdi
16	1320.5	48.5	648	3 Q9CIC6	Q9CIC6 caenorhabdi
17	1313	48.2	448	5 Q9GUA8	Q9GUA8 cryptospori
18	1307	48.0	448	5 Q9GUA8	Q9GUA8 cryptospori
19	1298	47.7	419	10 Q9M1N6	Q9M1N6 plasmodium

20	1259	46.3	566	3 Q9HDT2	Q9HDT2 usiliago ma
21	1240	45.6	437	5 Q9GUS9	Q9GUS9 cryptospori
22	1194	43.9	409	10 Q9PFA0	Q9PFA0 mus musculu
23	1138	41.8	429	5 Q9GUA8	Q9GUA8 cryptospori
24	1006	37.0	481	3 Q9PFA0	Q9PFA0 mus musculu
25	852	31.3	377	4 Q9PFA0	Q9PFA0 mus musculu
26	847	31.1	377	4 Q9PFA0	Q9PFA0 mus musculu
27	847	31.1	377	4 Q9PFA0	Q9PFA0 mus musculu
28	759.5	27.9	428	5 Q9GUA8	Q9GUA8 cryptospori
29	644.5	23.7	256	11 Q9PFA0	Q9PFA0 mus musculu
30	594.5	21.8	223	4 Q9PFA0	Q9PFA0 mus musculu
31	475	17.5	375	2 Q9PFA0	Q9PFA0 mus musculu
32	448.5	16.5	389	2 Q9PFA0	Q9PFA0 mus musculu
33	425.5	15.6	367	2 Q9PFA0	Q9PFA0 mus musculu
34	389.5	14.3	389	2 Q9PFA0	Q9PFA0 mus musculu
35	367.5	13.5	158	10 Q9LXN8	Q9LXN8 arabidopsis
36	340	12.5	883	5 Q9PFA0	Q9PFA0 mus musculu
37	317.5	11.7	380	2 Q9PFA0	Q9PFA0 mus musculu
38	309	11.4	1108	10 Q9PFA0	Q9PFA0 mus musculu
39	301.5	11.1	425	10 Q9PFA0	Q9PFA0 mus musculu
40	299.5	11.0	142	10 Q9PFA0	Q9PFA0 mus musculu
41	298.5	11.0	1215	4 Q9PFA0	Q9PFA0 mus musculu
42	297	10.9	310	2 Q9PFA0	Q9PFA0 mus musculu
43	296.5	10.9	1063	4 Q9PFA0	Q9PFA0 mus musculu
44	296.5	10.9	1066	4 Q9PFA0	Q9PFA0 mus musculu
45	293	10.8	577	10 Q9LXN8	Q9LXN8 arabidopsis

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	501 AA.
Q9SLZ3				
ID	Q9SLZ3			
AC	Q9SLZ3			
DT	01-MAY-2000 (TEMBLrel. 13, Created)			
DT	01-MAY-2000 (TEMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TEMBLrel. 17, Last annotation update)			
DE	HISTONE DEACETYLASE.			
GN	F20D10.250 OR ATG38130.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI-TaxID-3702;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Bevan M., Medler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,			
RA	Mayer K.F.X., Schueller C.;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RA	EU Arabidopsis sequencing project;			
RA	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RL	(3)			
RP	SEQUENCE FROM N.A.			
RA	Medler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,			
RA	Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	(4)			
RP	SEQUENCE FROM N.A.			
RA	EU Arabidopsis sequencing project;			
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RL	EMBL: AL03538; CAB37553.1; -			
DR	EMBL: AL161593; CAB80478.1; -			
DR	InterPro: IPR000286; His_deacetylase.			
DR	Pfam: PF00850; Hist_deacetyl; 1.			
DR	PRINTS: PR01270; HDASUPP.			
SQ	SEQUENCE 501 AA: 56023 MW; 857D2E3D16B7CC1F CRC64;			

Query Match 99.9%; Score 2719; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 5.3e-205;

Matches 500; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDTGNSLASGPDGVKRVKVCYFDPEVGNYYGCGHPMKPHRIKMTALLAHYGLLOHM 60
    |||||
Db 1 MDTGNSLASGPDGVKRVKVCYFDPEVGNYYGCGHPMKPHRIKMTALLAHYGLLOHM 60
QY 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
    |||||
Db 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
QY 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
    |||||
Db 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
QY 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
    |||||
Db 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
QY 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
    |||||
Db 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
QY 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
    |||||
Db 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
QY 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
    |||||
Db 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
QY 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
    |||||
Db 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
QY 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
    |||||
Db 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
QY 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
    |||||
Db 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
QY 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
    |||||
Db 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
QY 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
    |||||
Db 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
QY 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
    |||||
Db 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
QY 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
    |||||
Db 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
QY 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
    |||||
Db 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
QY 481 VKMEEGTNKGAEOAPPKT 501
    |||||
Db 481 VKMEEGTNKGAEOAPPKT 501
QY 481 VKMEEGTNKGAEOAPPKT 501
    |||||
Db 481 VKMEEGTNKGAEOAPPKT 501

```

```

Db 1 MDTGNSLASGPDGVKRVKVCYFDPEVGNYYGCGHPMKPHRIKMTALLAHYGLLOHM 60
QY 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
    |||||
Db 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
QY 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
    |||||
Db 61 VLKPPAREBRLCRFHADDDVSLRSTPTETODQIRQLKRFNVGDCPVFDGLYSFCOT 120
QY 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
    |||||
Db 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
QY 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
    |||||
Db 121 YAGSGVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 180
QY 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
    |||||
Db 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
QY 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
    |||||
Db 181 VDIDIHGGDVEEAFATDRVMTVSFHKFGDYFPGTGHIDIGYSGSKYSLNPLDDGI 240
QY 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
    |||||
Db 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
QY 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
    |||||
Db 241 DDESYSLLFKPIKGMKWEIFRPGAVVLCGADSLSDRLGCFNLSTKGHAECVFKMSFN 300
QY 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
    |||||
Db 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
QY 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
    |||||
Db 301 VPLLLOGGGYTTIRNVAKWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAAPSMEN 360
QY 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
    |||||
Db 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
QY 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
    |||||
Db 361 KNSRQMLEIRNDLLHNLSTLOHAPSVFQERPPDETPEVDEDOEDGDKRMDPSDMV 420
QY 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
    |||||
Db 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
QY 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
    |||||
Db 421 DDDKRPISRKREAVEPDTKDKGLKGMKGCCEVEDESGTAVTGVNPGVGEAS 480
QY 480 VKMEEGTNKGAEOAPPKT 499
    |||||
Db 480 VKMEEGTNKGAEOAPPKT 499
QY 480 VKMEEGTNKGAEOAPPKT 499
    |||||
Db 480 VKMEEGTNKGAEOAPPKT 499

```

Query Match 78.4%; Score 2134.5; DB 10; Length 500;

Query Match 72.0%; Score 1958.5; DB 10; Length 493;

Best Local Similarity 79.8%; Pred. No. 3.5e-159;

Best Local Similarity 77.5%; Pred. No. 2.2e-145;

Matches 399; Conservative 31; Mismatches 61; Indels 9; Gaps 3;

Matches 371; Conservative 36; Mismatches 59; Indels 13; Gaps 5;

QY 1 MDTGNSLASGPDGVKRVKVCYFDPEVGNYYGCGHPMKPHRIKMTALLAHYGLLOHM 60

QY 123 GGSVGSVYKLNHGLCDIAINMAGGLHNAKCEASGFCYVNDIVLAILLELKQHEBVL 182

Db 128 GSVGAVKLNHG-HDIAIMAGGLHNAKCEASGFCYVNDIVLAILLELKHQRLVYD 186
 QY 183 IDIHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDGID 242
 Db 187 IDIHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDGID 246
 QY 243 ESYHLEFKPIIMGKVMETFRPGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSFNV 302
 Db 247 DSVOSIFKPIISVMEYRFGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSFNV 306
 QY 303 LLLGGGGTIRIVNARCWCETGVALGVEEDKMPHEHYEYFGPDYTLHVASNMENK 362
 Db 307 LLLGGGGTIRIVNARCWCETGVALGHELDLDMKPNNEYEYFGPEYSLFVAASNMENR 366
 QY 363 SRONLEIRDLHNLSTKLOHAPSVPQERPPOTETPEVDEDDGDKRMDPSDMDVD 422
 Db 367 TKNOLEETKCNILDNLSKLOHAPSVOEERIPETKLPEDPEDDDPERIDDPDSM-VLD 425
 QY 423 DRKPIPS-----RVKRAVEPDTKDKDLKIGIMERGCEVEYDESGTKVTGVN 472
 Db 426 DKPMGHSASLLRNIEVKREITESAKDQG-KRLTTEHKGPEDPADDPDSSKQAPVS 483

RESULT 4
 ID Q9FML2 PRELIMINARY: PRT: 471 AA.
 AC Q9FML2:
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, last annotation update)
 DE HISTONE DEACETYLASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-98162728; PubMed-9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 Tabata S.;
 RT Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned pl clones.*;
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB008265; BAB1053.1; -;
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl. 1.
 DR PRINTS: PR01270; HDASUPER.
 DR SQUENCE 471 AA; 52651 MW; CAL6C2640D1B1732 CRC64;

Query Match 57.6%; Score 1568.5; DB 10; Length 471;
 Best Local Similarity 60.9%; Pred. No. 7.8e-115;
 Matches 293; Conservative 64; Mismatches 93; Indels 31; Gaps 5;

QY 2 DTGNSLASGPDGVKKKVCYFYDPEVGNYYGGGHPMKPHRIRMTALLAHYGLLOHMOV 61
 Db 4 DESGISLSPSPGDKRRKRVSYFEPTIGDYGGGHPMKPHRIRMAHSLIHHHLRLEI 63
 QY 62 LKPPAREDLRCFRHADYVSLRSTIPETQD--QIRLKRPNVGEDCVPFGDGLVSFCQ 119
 Db 64 SRSLDASDIGRHSPEYDVLASVSPESMGDPSAARNLRPNVGEDCVPFGDGLVSFCR 123
 QY 120 TVAGSGVSGVKLNHGLDIAIMMAGGLHNAKCEASGFCYVNDIVLAILLELKHQERVL 179
 Db 124 ASAGSGTGAAVKLNRODADIAIMMGGGLHNAKCEASGFCYVNDIVLAILLELKHQERVL 183
 QY 180 YVIDIDHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDG 239
 Db 184 YVIDIDHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDG 243

QY 240 IDDESYLEFKPIIMGKVMETFRPGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSF 299
 Db 244 MDEDSRSLFRPLQKVMETFRPGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSF 303
 QY 300 NVPLLLGGGGTIRIVNARCWCETGVALGVEEDKMPHEHYEYFGPDYTLHVASNMENK 359
 Db 304 NVPLVGGGGTIRIVNARCWCETGVALGVEEDKMPHEHYEYFGPDYTLHVASNMENK 363
 QY 360 NKNRQMLEIRDLHNLSTKLOHAPSVPQERPPOTETPEVDEDDGDKRMDP----- 414
 Db 364 NLPTRQMERIRNTLEQLSLHAPSVOFHTPPNVR--LDEPEDMETPRKPIWSG 421
 QY 415 --DSMDVDKRPISRVKRAVEPDTKDKDLKIGIMERGCEVEYDESGTKVTGVN 472
 Db 422 TATYEDSDDDDKP-----LHGYSGRG-GATTDRSTGDEMDDN 461
 QY 473 P 473
 Db 462 P 462

RESULT 5
 ID Q9FVE5 PRELIMINARY: PRT: 471 AA.
 AC Q9FVE5:
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, last annotation update)
 DE PUTATIVE HISTONE DEACETYLASE.
 GN RPD3B.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Wu K., Malik K., Tian L., Brown D., Miki B.;
 RT Functional analysis of RPD3 histone deacetylase homologs in
 RT Arabidopsis thaliana.*;
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF195548; AAG28475.1; -;
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl. 1.
 DR PRINTS: PR01270; HDASUPER.
 DR SQUENCE 471 AA; 52720 MW; 371BF7040E508849 CRC64;

Query Match 57.4%; Score 1561.5; DB 10; Length 471;
 Best Local Similarity 60.7%; Pred. No. 2.8e-114;
 Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

QY 2 DTGNSLASGPDGVKKKVCYFYDPEVGNYYGGGHPMKPHRIRMTALLAHYGLLOHMOV 61
 Db 4 DESGISLSPSPGDKRRKRVSYFEPTIGDYGGGHPMKPHRIRMAHSLIHHHLRLEI 63
 QY 62 LKPPAREDLRCFRHADYVSLRSTIPETQD--QIRLKRPNVGEDCVPFGDGLVSFCQ 119
 Db 64 SRSLDASDIGRHSPEYDVLASVSPESMGDPSAARNLRPNVGEDCVPFGDGLVSFCR 123
 QY 120 TVAGSGVSGVKLNHGLDIAIMMAGGLHNAKCEASGFCYVNDIVLAILLELKHQERVL 179
 Db 124 ASAGSGTGAAVKLNRODADIAIMMGGGLHNAKCEASGFCYVNDIVLAILLELKHQERVL 183
 QY 180 YVIDIDHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDG 239
 Db 184 YVIDIDHGGGVEAEATYADRVMTVSFHKFGDFPGTGHIDIGSGSKYSLNVLDDG 243
 QY 240 IDDESYLEFKPIIMGKVMETFRPGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSF 299
 Db 244 MDEDSRSLFRPLQKVMETFRPGAVVLOCGADSLSGDLGCFNLSTIKGHAECVFMRSF 303

Query	Match	56.9%	Score 1548	DB 10	Length 458
300	NPVLLTLLGGGGTIRNVAVCMCEYGVGALVEEDKMPHEVEYFGPDTLHVAPSNME	Best Local Similarity 63.3%	Pred. No. 3e-113		
304	NVPPLVAVLGEGGTTIRNVAVCMCEYGVGALVEEDKMPHEVEYFGPDTLHVAPSNME	Matches 280	Conservative 66	Mismatches 80	Indels 16
360	NKNSKQMLEEIRNDLHLNLSKLOHAPSVPEOERPPDTEPEVEDQEDGDKRMDFDS---				
364	NLNTPKDKMERIRNTLLEQSLIHPASVQFQTPPVNRY--LDEDEDDETRPKPRXMSG				
417	----DMDVDDDKRPILPSRYKREAVBPDTKDKDGLGIMERGKGEVEVDESGSTVYGTGN				
422	TATVESDSDDDKP-----LHGYSGRG--GATYTDRTSTGEDEMDDON				
473	P 473				
462	P 462				
09ZTP8	09ZTP8	PRELIMINARY	PRT	458 AA	
AC	09ZTP8				
DT	01-MAY-1999 (TREMBlrel. 10, Created)				
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)				
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)				
DE	HISTONE DEACETYLASE.				
GN	HDB.				
OS	Zea mays (Maize).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;				
OC	Panicoidae; Andropogoneae; Zea.				
OX	NCBI_TaxId:4577;				
RN	11				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-CV. CUZCO.				
RA	Pipal A., Wegener S.;				
RL	Submitted (Jan-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF045473; AAD10139.1; -				
DR	Mendel: 39025; Zeama:3043;39025.				
DR	InterPro: IPR000286; His_deacetylase.				
DR	Pfam: PF00850; Hist_deacetyl1; 1.				
DR	PRINTS: PR01270; HDASUPER.				
DR	SEQUENCE 458 AA; 50940 MW; A62775068225B679 CRC64;				
Query Match	56.9%	Score 1548	DB 10	Length 458	
Best Local Similarity	63.3%	Pred. No. 3e-113			
Matches 280	Conservative 66	Mismatches 80	Indels 16	Gaps	
07	7 SLASRPDGVKRVVCFYDPEGVNYYGCGHPKPRIRIKTHALIAHYGLQIMOVYIKRPP				
DB	12 SPAGGDAHRAVRVSFYEPSIDYTGGOCHPKPRIRIRNAHSLVHYGHLRLLELSRPP				
07	67 AREERDLCRFHADYVSLRSITPETQQDIPQLKRFNVGEGDPVFDGLYSCOTYAGGSV				
DB	72 ASEDRIRRHSDYAFGLASATGNGVLDPRAIKRVNVEGDEPVDGLPFPQASAGSGI				
07	127 GGSVKLNLHGCDIAINMAGGLHRAKCEASGFCYVNDIVLALILELKKQHERVLYVDIH				
DB	133 GAAYVILNRGDADITVMAGGLHRAKCEASGFCYVNDIVLALILELKKFRRVLYVDIVH				
07	187 HEDGVFEAFYADRVNTVSFHKFGDYFTPGTHIDIGYSSGYYSLNVLVDGIDDESH				
DB	192 HEDGVFEAFYADRVNTVSFHKFGDYFTPGTHIDIGYSSGYYSLNVLVDGIDDESH				
07	247 LLEKPIGKVMELIFRGAVALVLOGGADSLSGDRIGCFNLSIKGHAEVYKPMRSFNPLLL				
DB	252 GLFQCLIKVMYVYQPDVYVLOGGADSLSGDRIGCFNLSIKGHAEVYKPMRSFNPLLL				
07	307 GGGGTYTIRNVAVCMCEYGVGALVEEDKMPHEVEYFGPDTLHVAPSNMEKNSRQ				
DB	312 GGGGTYTIRNVAVCMCEYGVGALVEEDKMPHEVEYFGPDTLHVAPSNMEKNSRQ				
07	367 LEEIRNDLHLNLSKLOHAPSVPEOERPPDTEPEVEDQEDGDKRMDFDSMDVDDDKP				

[illegible]

09VZAI	09VZAI	PRELIMINARY;	PRT;	521 AA.
ID	09VZAI			
AC	09VZAI;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	RPD3 PROTEIN.			
GN	RPD3 OR CG7471.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BERKELEY.			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Amadis M.D., Celniker S.E., Holt R.A., Evans C.A., Goebye J.D.,			
RA	Amadis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abbil J.F., Aghayani A., An H.-J., Andrews-Piankoff C., Baldwin D.,			
RA	Ballem R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Betman B.P., Bhanderi D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,			
RA	Gloeder A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,			
RA	Jatani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Laoko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-T., Wasserman D.A., Weinstock G.M., Weissbach J.,			
RA	Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yah R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,			
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,			
RT	"The genome sequence of Drosophila melanogaster.";			
RL	Science 287:2185-2195(2000).			
DR	EMBL; AE003482; AAF47924.1; -			
DR	flybase; FBgn0015805; Rpd3			
DR	Interpro; IPR000286; His_deacetylase.			
DR	Pfam; PF00850; Hist_deacetyl_1.			
DR	PRINTS; PR01270; HDASUPER			
50	SEQUENCE 521 AA; 56330 MW; B0F6503D42A1BC69 CRC64;			

[illegible]

Db 66 HSD5EVRFLRSIRPDNNSEYKNKQMFNFVNGEDCPVEDGILYFCQJLSAGGSVAANAALKKQ 125

QY 136 LQDIAINMAGGLIHAHAKKKEASGFCYVNDIVYALILELLKQHEVLYVDIDIHGGDYEAFA 195

Db 126 ASELCTINMGGLIHAHAKKKEASGFCYVNDIVYALILELLYHQRVLYTIDIVHGGDYEAFA 185

QY 196 YATDRVMYVSEHKFGDYPPGTGHIODIGYSGSKYSLNVPIDDDIDDSHYLLFFPKIMK 255

Db 186 YTTDRVMYVSEHKFGDYPPGTGDLRIDIGAGKRYAAVMIPLRDGDDAEIEFPILISK 245

QY 256 VMEIFRQCAVYVLOGGADSLSGDRLGCFNLSTKGAAECYKFMKSFVPLLLGGGYTIRN 315

Db 246 VMEIFQPAVAVVLOGGADSLTGDRILGCFNLTYKGHGKCYEFYKYNLPELWMGGGGYTIRN 305

QY 316 VARQWCYETGALCYEVDKMDPEHETRYPGPDYLLIHAAPSMEKNSROMLEETIRNL 375

Db 306 VSRCTIYETSVALLAVEINELPLNYDFEYFGPDRKLHISPMIMNTSEYLEKIKNRLF 365

QY 376 HNLKSLQIHAAPSVPQERPPDETPEVEDD-----QEDGDKRMPDSD-MDVDD-- 422

Db 366 ENLNLPIHAPEVQIQAIPEDAINDESDQEDRYDKDRLPQSDKDKRIYENESYSEDEG 425

QY 423 -----DKRPISRYKREAVPEPTADKDG-----LKLIMEKGCCYEVDSEGST 466

Db 426 EGRGRDNRSYVGQRKRPRLDKDTNSNKASSETSEIKDEKEKGADGDEESTASNT 481

RESULT 9

Q99PA0 PRELIMINARY: PRI: 428 AA.

AC Q99PA0; DT 01-JUN-2001 (TREMblrel, 17, Created)

DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)

DT 01-JUN-2001 (TREMblrel, 17, Last annotation update)

DE HISTONE DEACETYLASE 3.

GN HDAC3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=WTSTAR; TISSUE=TESTIS;

RA Milquet V., Chavez M., Korbets R., Geerts A.;

RT "Expression pattern of rat histone deacetylases";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF321131; AAK11184.1; -

SO SEQUENCE 428 AA; 48829 MW; 7CB4480CFE33041FE CRC64;

Query Match 50.58; Score 1374; DB 11; Length 428;

Best Local Similarity 56.68; Pred. No. 1.2e-99;

Matches 241; Conservative 73; Mismatches 90; Indels 22; Gaps 3;

QY 17 RKVCYFVDEYGNYYGGGHPMKPHRIMKTHALLAHYGLDLMQVLYKPPARERDLGRFH 76

Db 3 KTVAFYDFDPDVGNGNHYGAGHPMKPHRLALHTSLVHYLYKMYFKFYQASQHDCKRFH 62

QY 77 ADDYVSELRSTIPETQDQIRQLKRFNVGEDCPVEDGLYFCQJYAGGSVGSYVNLHGL 136

Db 63 SEDYIDFLQVRSPINMGQFTSLNLFNVGDDCPVPPGLFECFSRYTASLGATGATLNKKI 122

QY 137 CDIAINMAGGLIHAHAKKKEASGFCYVNDIVYALILELLKQHEVLYVDIDIHGGDYEAFA 196

Db 123 CDIAINMAGGLIHAHAKKKEASGFCYVNDIVYALILELLKQHEVLYVDIDIHGGDYEAFA 182

QY 197 ATDRVMYVSEHKFGDY-FPGTGHIDDIGYSGSKYSLNVPIDDDIDDSHYLLFFPKIMK 255

Db 183 LTDRVMYVSEHKFGDYPPGTGDMTEVGAESGRYCLNVPRLRDGIDQGSYKRLFPVISO 242

QY 256 VMEIFRQCAVYVLOGGADSLSGDRLGCFNLSTKGAAECYKFMKSFVPLLLGGGYTIRN 315

Db 243 VMEIFRQCAVYVLOGGADSLSGDRLGCFNLSTKGAAECYKFMKSFVPLLLGGGYTIRN 302

	Query Match	50.5%	Score 1374	DB 11	Length 428
	Best Local Similarity	56.6%	Pred. No. 1,2e-99		
Matches	241	Conservative	73	Mismatches	90
				Indels	22
				Gaps	3
QY	17	RKVCYFDEEVGNTYTGCGHPMKPHRIIMTHALLAHYCLLOHMOVLKPPAREKOLCRFH	76		
DB	3	KTVAYFDEPDGNGHYHGHGHPMKPHRIALITHSLVLYHLYGUKRMITFKFYQASQHMCRFH	62		
QY	77	ADYVSEFLRSTPPEOQOUIROLKRFENYGECPYVDGYLSTFCQTAGSGVGSYKLNGL	136		
DB	63	SEDIYIDFLQRRSPINMGQFTKSLANFVNGDDCPYRPGLFECSTRTGSLGATOLDANKKI	122		
QY	137	CDIAINMAGGIHHAHKCCASGFCYVNDIVLAILELLOHERVLYVDIDIIHGGVEEAFY	196		
DB	123	CDIAINMAGGIHHAHKCFEASGFCYVNDIVLILLELTKHPHLYLIDIDIIHGGVOEAFY	182		
QY	197	ATDQVMTVSEFHKFDY-FPTGHIODIGYSGSKYSLWLPDDGDIDESYHLFKPIAGK	255		
DB	183	LTDQVMTVSEFHKQNYFFPPTGDMTEVGAESGRYTCIPLVRLDGDIDDSYKHLRQPIVISO	242		
QY	256	VMELFRCAVYVLOGGADSLSDRLGCFNLSIKGAECYKFMKSFNVPIILLGGGYTIRN	315		
DB	243	VVDYEQPCYVLOGGASDLGCDRLGCFNLSIRGHECEYKSFNPIPLTVLGGGGYTIRN	302		

```

QY 316 VARWCYETGVALGEVEDKMPHEHYEFGPDYTLHV-APSNNKNSKQMLEIRNDL 374
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 303 VARCWYETSLVEEALISEELPYSEFEYFAPDPTLHPDVSTRIENQNSROYLDQIROT 362

QY 375 LHNLSKLOHAPSVFQERPPDTE-----PEVDEQEOGDKRMDP 414
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 363 FENLKMNLNHAISVQIHVPADLLTYORTDEADAEERGPENYSRPAENFEYDGDHNDK 422

QY 415 DSDMDV 420
| | | | | : : : : :
Db 423 ESDVEI 428

RESULT 10
Q9P4F5 PRELIMINARY: PRT: 687 AA.
AC 09P4F5:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HISTONE DEACETYLASE RPD3A.
GN RPD3A.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461766; PubMed=11004483;
RA Graessle S., Dangl M., Haas H., Malt K., Trojer P., Brandtner E.M.,
RA Walton J.D., Loidl P., Brosch G.,
RT Characterization of two putative histone deacetylase genes from
RT Aspergillus nidulans.
RL Blochim. Biophys. Acta 1492:120-126(2000).
DR EMBL: AF163862; AAF80489.1;
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl.
DR PRINTS: PR01270; HDASUPER.
SO SEQUENCE 687 AA; 75431 MW; CFC7C6535236451 CRC64;

Query Match 50.08; Score 1361.5; DB 3; Length 687;
Best Local Similarity 49.28; Pred. No. 2.3e-98;
Matches 261; Conservative 89; Mismatches 136; Indels 45; Gaps 9;

QY 10 SGRPGV-----KAKVCYFDPDEVNYYGCGHPRKPHRIKTHALLAHYGL 56
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 6 SGRAGPPLDPLDLNVSDRSKRVAFYDSQVGNAYVSGHMKPHRIKTHSLVANNYSLY 65

QY 57 QHGVLPKRFARERDLCRFNADYVSFLRSITPETQODIROLKRFNVEDCDVFDGLYS 116
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 66 KKEIYRAKRPASKFEMTOFHTDEYIDFLSKVTPDNMDAFKEDSKYVNGDCCPVFDLFE 125

QY 117 FCGYTAGSGVSGSKLHNGLCDIAINNAGGLHAKKCEASGFCVNDIVLALLELKH 176
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 126 FCGISAGSGMGAARLNKNCODIANNAGGLHAKKSEASGFCVNDIVLALLELKH 185

QY 177 RVLVVDIDIHGDCVEAFATDVRMTVSFHKFEGDYFGTGHIDIDIGSGKYSLVNPL 236
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 186 RVLVVDIDVHNGDVEAFYTTDRVMYVSFHKYGEYFGTGHIDIGVGCKYAAVNPL 245

QY 237 DQGITDESAYHLKFRIMGKWEIFRGAVALLOCADSLSDORLCPNLISKGHAECVFM 296
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 246 RDGITDESAYKSIFFYKISVYEMVRPRAVVLQCGDLSLSDORLCPNLSMKGHAECV 305

QY 297 RSEFVPLLLLOGGGYTTIRNARWCYETGVALGEVEDKMPHEHYEFGPDYTLHVAPS 356
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 306 KSEFVPLLLVGGGGYTKRANVARTWAFEGILVGNLSELPLNDYEFYFAPDELVDVPS 365

QY 357 NMEKNSKQMLEIRNDLHNLISKLOHAPSVFQERPPDTEPEVDEQEOGDKRMDPDS 416
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 366 NMDANRREYLDKIRTOYVENMLKRTAFAPSVQMTDVFRE---PLVDGMDDEAFALD-DL 421

```

```

QY 417 DMDVDDRRKPIPSRVKREAVP-----DTKDKD--GLKGI-----MERKKGCEVEVD 461
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 422 DEDENKDKRRTKRRPQYVEKPEGLSDSEDEDENAANGVTRKPAHLKRRNOANYRLDAD 481

QY 462 ---ESGSTKATGVNPGVEE---ASYKM-----EEEGTKGAGCAQAPR 499
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 482 SGVESGMATPDQASSVADEEMDTGTDVKTETAPPEPDSBAQTSAAEP 532

RESULT 11
Q9JMO8 PRELIMINARY: PRT: 428 AA.
AC 09JMO8:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HISTONE DEACETYLASE 3.
GN HDAC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J;
RX MEDLINE=20391214; PubMed=10542131;
RA Dangond F., Foerzler D., Meremowicz S., Morton C.C., Belier D.R.,
RA Gullans S.R.,
RT Cloning and expression of a murine histone deacetylase 3 (hdac3)
RT cDNA and mapping to a region of conserved syteny between murine
RT chromosome 18 and human chromosome 5.
RL Mol. Cell Biol. Res. Commun. 2:91-96(1999).
DR EMBL: AF098295; AAF28798.1;
DR MGD: MGI:1343091; Hdac3.
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl.
DR PRINTS: PR01270; HDASUPER.
SO SEQUENCE 428 AA; 48890 MW; 995EFB827568804 CRC64;

Query Match 49.98; Score 1359; DB 11; Length 428;
Best Local Similarity 56.18; Pred. No. 1.8e-98;
Matches 239; Conservative 73; Mismatches 92; Indels 22; Gaps 3;

QY 17 RKYCYFDPDEVNYYGCGHPRKPHRIKTHALLAHYGLDHOVYLPKPARERDLGRPH 76
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 3 KYVAFYDPLDVGPNHPRGAGHPRKPHRLALHSLVLYHGYLKKMIVKPYQASQHDRCRH 62

QY 77 ADDYVSFLRSITPETQODIROLKRFNVEDCDVFDGLYSFCOTYAGSVGSKVNLHGL 136
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 63 SEDYIDFLQVSPJTNMGCFKSLDAFNVGDDCPVFGLEFCGRYTGASLQCATQIANNK 122

QY 137 CDIAINNAGGLHAKKCEASGFCVNDIVLALLELKHNRVLYVVDIDIHGDCVEAFY 196
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 123 CDIAINNAGGLHAKKCEASGFCVNDIVLALLELKHNRVLYVVDIDIHGDCVEAFY 182

QY 197 ATDRVMTVSFHKFGDY-FPGTGHIDIDIGVSGKYSLVNPLDQIDDESAYHLFKPMK 255
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 183 LTRVMTVSFHKGNFFFGADMYEAGSGRYUCLNVPLRGIDIDQYKHLFGVVISQ 242

QY 256 VMEIFRGAVALLOCADSLSDORLCPNLISKGHAECVFMKSFVNPLLLGGGGYTTIRN 315
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 243 VVDFYPTCIVLOCADSLSDORLCPNLISIRGHGCEVEYKSFNPLLLVGGGGYTVRN 302

QY 316 VARWCYETGVALGEVEDKMPHEHYEFGPDYTLHV-APSNNKNSKQMLEIRNDL 374
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 303 VARCWYETSLVEEALISEELPYSEFEYFAPDPTLHPDVSTRIENQNSROYLDQIROT 362

QY 375 LHNLSKLOHAPSVFQERPPDTE-----PEVDEQEOGDKRMDP 414
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 363 FENLKMNLNHAISVQIHVPADLLTYORTDEADAEERGPENYSRPAENFEYDGDHNDK 422

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:17 ; Search time 17.07 seconds

(without alignments)
1076.103 Million cell updates/sec

Title: US-09-645-337-2

Perfect score: 2722

Sequence: 1 MDTGNSLASGPDGVKRYVC.....KMEEGTNGAGAPFPT 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2722	100.0	501	1 HDAC_ARATH	022446 arabidopsis
2	1986.5	73.0	513	1 HDAC_ARATH	P65521 zea mays (m
3	1554	57.1	482	1 HDAL_MOUSE	009106 mus musculu
4	1544	56.7	482	1 HDAL_MOUSE	013547 homo sapien
5	1538	55.5	480	1 HDAL_CHICK	P65517 gallus gall
6	1521	55.9	480	1 HD12_XENLA	042227 xenopus lae
7	1520	55.8	480	1 HD11_XENLA	P31695 xenopus lae
8	1513.5	55.6	488	1 HDAL_MOUSE	P70288 mus musculu
9	1507	55.4	576	1 HDAL_STRPU	P65518 strongyloce
10	1494	54.9	488	1 HDAL_HUMAN	092769 homo sapien
11	1478	54.3	520	1 HDAC_DROME	094517 drosophila
12	1465.5	53.8	488	1 HDAL_CHICK	P65519 gallus gall
13	1392	51.1	461	1 HDAL_MOUSE	017695 caenorhabdi
14	1376	50.6	428	1 HDAL_MOUSE	P65520 gallus gall
15	1373	50.4	428	1 HDAL_MOUSE	015379 homo sapien
16	1366	50.2	424	1 HDAL_MOUSE	088895 mus musculu
17	1341	49.3	433	1 RPD3_YEAST	P32561 saccharomyc
18	1167.5	42.9	434	1 PHD1_SCHPO	013298 schizosacch
19	1077	39.6	507	1 HDAL_MOUSE	009440 caenorhabdi
20	1066.5	39.2	452	1 HOS1_YEAST	P33096 saccharomyc
21	471.5	17.3	470	1 HOS1_YEAST	012214 saccharomyc
22	443.5	16.3	387	1 ACUC_BACSU	P39067 bacillus su
23	386.5	14.2	385	1 ACUC_STAXY	056195 staphylococ
24	352	12.9	359	1 Y130_ARCFU	030107 atcheaoglob
25	320.5	11.8	706	1 HDAL_MOUSE	P33973 saccharomyc
26	298.5	11.0	1215	1 HDAL_MOUSE	094977 homo sapien
27	278	10.2	687	1 HDAL_MOUSE	P65523 schizosacch
28	275.5	10.1	1080	1 HDAL_MOUSE	P33038 gallus gall
29	275.5	10.1	1149	1 HDAL_MOUSE	092245 mus musculu
30	274	10.1	1113	1 HDAL_MOUSE	092246 mus musculu
31	266.5	9.8	1122	1 HDAL_MOUSE	094977 homo sapien
32	263	9.7	1084	1 HDAL_MOUSE	P65524 schizosacch
33	261.5	9.6	310	1 YGLA_SYNP2	P28606 synecococc

ALIGNMENTS

RESULT	1	STANDARD	PRT	501 AA
HDAC_ARATH				
AC	022446;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	HISTONE DEACETYLASE (HD).			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi			
OX	NCBI_TaxID:3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN:CV. COLUMBIA;			
RA	Tomihama T., Shoji K., Hanyu H., Okano T.;			
RL	Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON			
CC	THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).			
CC	HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL			
CC	REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY			
CC	SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA			
CC	FAMILY. HD SUBFAMILY 1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF014824; AAB6486.1; -			
DR	InterPro; IPR000286; His_deacetylase.			
DR	Pfam; PF00850; Hist_deacetyl; 1.			
DR	PRINTS; PR01270; HDASUPER.			
DR	PRINTS; PR01271; HISDACETYLASE.			
KW	Hydroxylase; Nuclear protein.			
SQ	SEQUENCE 501 AA; 56037 MW; C50AF5624958D6C2 CRC64;			

Query Match Score 2722; DR 1; Length 501;

Best Local Similarity 100.0%; Pred. No. 1.2e+203;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDTGNSLASGPDGVKRYVCYFDPEVGNYYGCGHMKPHRIMTHALLAHYGLDLMHQ	60
DB	1	MDTGNSLASGPDGVKRYVCYFDPEVGNYYGCGHMKPHRIMTHALLAHYGLDLMHQ	60
QY	61	VLAFFPARERDLRFHADDVYSLRSTPPTGQDQIQQLKRFNVGECPPVFDGLYSFCOT	120
DB	61	VLAFFPARERDLRFHADDVYSLRSTPPTGQDQIQQLKRFNVGECPPVFDGLYSFCOT	120

QY 121 YAGSGVGSVKLNHGLCDIAIMAGGLHAKKCEASGFCYNDIVLALIELLKHERVLY 180
 Db 121 YAGSGVGSVKLNHGLCDIAIMAGGLHAKKCEASGFCYNDIVLALIELLKHERVLY 180
 QY 181 VDIDHHDGVEAFATDRMTVSFHKFGDYFPGTGIDIGGSGKYSLANPLDDGI 240
 Db 181 VDIDHHDGVEAFATDRMTVSFHKFGDYFPGTGIDIGGSGKYSLANPLDDGI 240
 QY 241 DDESYHLFKPIMCKVMEIFRPGAVVLCGADSLSGDLGCFNLISIKHACVFMRSFN 300
 Db 241 DDESYHLFKPIMCKVMEIFRPGAVVLCGADSLSGDLGCFNLISIKHACVFMRSFN 300
 QY 301 VPLLLGGGGTTIRNVACWCYETGVALGVEEDKMPHEHYETFGPDYTLHVAPSNMEN 360
 Db 301 VPLLLGGGGTTIRNVACWCYETGVALGVEEDKMPHEHYETFGPDYTLHVAPSNMEN 360
 QY 361 KNSRMLBEIRNDLLHNLSKLOHAPSVFQERPPDETPEYDEDEDDKHPDSDMDV 420
 Db 361 KNSRMLBEIRNDLLHNLSKLOHAPSVFQERPPDETPEYDEDEDDKHPDSDMDV 420
 QY 421 DDDRPFPSRVKREAVEPDTKDKGLKIMRGKCEVEVDESGSTKVTGVNPGVEEAS 480
 Db 421 DDDRPFPSRVKREAVEPDTKDKGLKIMRGKCEVEVDESGSTKVTGVNPGVEEAS 480
 QY 481 VKMEEGTNGKGAQAFPPKT 501
 Db 481 VKMEEGTNGKGAQAFPPKT 501

RESULT 2
 HDAL_MOUSE STANDARD: PRT: 513 AA.
 AC P56521;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE HISTONE DEACETYLASE (RPO3 HOMOLOG).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 ON NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. W22;
 RA Rossi V., Hartings H., Motto M.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / ALPHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF035815; AAC50038.1;
 CC InterPro: IPR000286; His_deacetylase.
 CC Pfam: PF00850; Hist_deacetyl_1.
 CC PRINTS: PR01270; HDASUPER.
 CC PRINTS: PR01271; HISDACETLASE.
 CC Hydrolase: Nuclear protein.
 CC SEQUENCE 513 AA; 57546 MW; C45387CF3A38906F CRC64;

Query Match 73.0%; Score 1986.5; DB 1; Length 513;
 Best Local Similarity 75.2%; Pred. No. 1,3e-146;
 Matches 377; Conservative 44; Mismatches 61; Indels 19; Gaps 7;

QY 3 TGGNSLAS-PPDGVKRRVCFYDPENGVYGGGHPKPRIRIMTHALLAHYGLILDMOV 61
 Db 8 SGGNSLSVGGDGGKRRVCFYDPDVGNYTGGGHPKPRIRIMTHALLAHYGLILDMOV 67
 QY 62 LKPPARERDLCPHADYVSFLRSITPETQDQIRQLKPNVNGEDCPVEFDGYSFCQY 121
 Db 68 YRPAPARERLCPHAEYINFLRSVTPETQDQIRQLKPNVNGEDCPVEFDGYSFCQY 127
 QY 122 AGSGVGSVKLNHGLCDIAIMAGGLHAKKCEASGFCYNDIVLALIELLKHERVLY 181
 Db 128 AGASVGAVFNNHG-HDIAIMSGGLHAKKCEASGFCYNDIVLALIELLKHERVLY 186
 QY 182 DIDHHDGVEAFATDRMTVSFHKFGDYFPGTGIDIGGSGKYSLANPLDDGI 241
 Db 187 DIDHHDGVEAFATDRMTVSFHKFGDYFPGTGIDIGGSGKYSLANPLDDGI 246
 QY 242 DDESYHLFKPIMCKVMEIFRPGAVVLCGADSLSGDLGCFNLISIKHACVFMRSFN 301
 Db 247 DESYSLFKPIMCKVMEIFRPGAVVLCGADSLSGDLGCFNLISIKHACVFMRSFN 306
 QY 302 PLLLLGGGGTTIRNVACWCYETGVALGVEEDKMPHEHYETFGPDYTLHVAPSNMEN 361
 Db 307 PLLLLGGGGTTIRNVACWCYETGVALGVEEDKMPHEHYETFGPDYTLHVAPSNMEN 366
 QY 362 NSRMLBEIRNDLLHNLSKLOHAPSVFQERPPDETPEYDEDEDDKHPDSDMDV 421
 Db 367 NTRQQLDDIRS-KLSKLHAPSVFQERPPDETPEYDEDEDDKHPDSDMDV 422
 QY 422 DD-----RKPIPSRVKREAVEPDTKDKGLKIMRGKCEVEVDESGSTKVTGVNPGVEEAS 480
 Db 423 DHKAVESSRSRSLIGIKIKREFGENATRYODGGR-VASERHGLEPMEDIGSSKQAFQAD 481
 QY 471 VNPYGVVEAS-VKMEEGTNGK 490
 Db 482 ASAMAIIDPSVKNKPEPSTK 502

RESULT 3
 HDAL_MOUSE STANDARD: PRT: 482 AA.
 ID HDAL_MOUSE
 AC 009106; P97476;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HISTONE DEACETYLASE 1 (HD1).
 GN HDAC1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=97415582; Pubmed=9271381;
 RA Bartl S., Taplick J., Lager G., Khler H., Kuchler K., Seiser C.;
 RT Identification of mouse histone deacetylase 1 as a growth factor-
 RT inducible gene. J. Biol. Biol. 17:5033-5043(1997).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Johnson C.A.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY).
 CC -1- FUNCTION: ACTS AS A TRANSCRIPTIONAL REPRESSOR.
 CC -1- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY). ALSO FORMS A

```

CC      COMPLEX WITH SIN3 AND SAP18 (BY SIMILARITY).
CC
CC      -1- SUBCELLULAR LOCATION: NUCLEAR.
CC      -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHER LEVELS IN THYMUS
CC      AND TESTIS AND LOWER LEVELS IN LIVER.
CC      -1- INDUCTION: BY INTERLEUKIN-2.
CC      -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC      FAMILY. HD SUBFAMILY 1.
CC-----
CC      THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
CC      AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC      ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb.ch/announce/).
CC      OR SEND AN EMAIL TO license@isb-sib.ch).
CC
CC      EMBL: X98207; CAA66870.1; -.
CC      EMBL: U80780; AAB68398.1; -.
CC      MGD: MGI:108086; Hdac1.
CC      InterPro: IPR000286; His_deacetylase.
CC      Pfam: PF00850; Hist_deacetyl; 1.
CC      PRINTS: PRO1270; HDASUPER.
CC      PRINTS: PRO1271; HSDACETLASE.
CC      Hydrolase: Nuclear protein.
CC
CC      SEQUENCE 462 AA; 55075 MW; 7664D3C17F5E4844 CRC64;

```

Query Match	57.1%	Score 1554	DB 1	Length 482
Best Local Similarity	60.0%	Pred. No. 4.3e-113		
Matches 294	Conservative 64	Mismatches 98	Indels 34	Gaps 7

QY	14	GVKRVVCYFDPENGVNYUYGOGNHPMKPHIRMTAHLALLAHYGLLOHMQVLLKPPAREIDLC	73
Db	6	GTRKRVVCYUYDGVGNYUYGOGNHPMKPHIRMTHTHLLNTLNYGLYRMETIRPHKAAEMET	65
QY	74	RPHADIVSFLRSITPETOQDQIRQLKRPNGEDCPRVDGLYSFCQYAGSVGSGVKIN	133
Db	66	KYHSDDYIKFLRSIRPDNMSEYKQMOFRNVEGDCPRVDFGLJFEFCOLSTGGSVASAVKLN	125
QY	134	HGLDIAINMAGGLHNHAKKCEASGFCYVNDIYALATLELLKOHENRUYVDIHHDDGYEE	193
Db	126	KQODDIAANMAGGLHNHAKKSEASGFCYVNDIYALATLELLKTHQRFVLYIDIDHHGGVVE	185
QY	194	AFYATDRAVTVSFKHFGDYFPQTCGHIODIGVSGKYYUSLNPBLDDGIDDESHTLLFKPTM	253
Db	186	AFYTTDRAVTVSFKHGYEYFPCTGDLRDIAGAKGYAYANVPLRBDIDDESHTLAFKPYM	245
QY	254	GKVMETIFRPGAVYLOCGADSLSDGLDLCFNLSIKGHACVYKMRSENPVLLLGGGGYIT	313
Db	246	SKVMMFOPRSAVYLOCGSDSLSDGLDLCFNLTIKGHAKCVBFVKSPNLPMLLGGGGYIT	305
QY	314	RNVACWCYECGTVALGVENEDEKPRHEHYEGVGPYUHLVHVASNNENKSRQMLEIRND	373
Db	306	RNVACWYETIETVALDTEIPNELPNDYFEXFGPDKLISPSNMNTNLTNLEYLEXIKOR	365
QY	374	LLHNLSKLOHAPSVPEOERPDTEPREV-DEPOEDGDKRMD-----PDSDM	418
Db	366	LFENLRMLRPHACGYOMQALPREDALIEESDDEBEDRDKRISICSSDKRLACEEFSDSOE	425
QY	419	DVDDDKRPIPS-----RYKREAVBDPTDKDQSLKGMERGKCCVEYEDBSGSTKYTVGNP	473
Db	426	EEGEGKRKNSNFKKAKRVYTE---DEKRD-----PEEK-----EVTDEKTKKEKPDA	472
QY	474	VGVEEASAYKM	483
Db	473	KGVK-EVKL	481

RESULT	4
HDAL_HUMAN	
ID	HDAL_HUMAN
AC	Q13547; Q92534;
DT	01-NOV-1997 (Rel. 35, Created)
STANDARD;	PRT;
482 AA	

```

DF 01-NOV-1997 (Rel. 35, last sequence update)
DT 20-AUG-2001 (Rel. 40, last annotation update)
DE HISTONE DEACETYLASE 1 (HD1).
GN HDAC1 OR RPD3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RX NCBI_TaxId=9606;
RX [1]
RP SEQUENCE FROM N.A.
RP TISSUE=T-cell;
RX MEDLINE=96185499; PubMed=8602529;
RA Taunton J., Hassig C.A., Schreiber S.L.;
RT "A mammalian histone deacetylase related to the yeast transcriptional
RT regulator Rpd3p." ;
RL Science 272:408-411(1996).
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Fetal Lung;
RX MEDLINE=96244606; PubMed=8646880;
RA Furukawa Y., Nakakami T., Sudo K., Inazawa J., Matsumine A.,
RA Akiyama T., Nakamura Y.;
RT "Isolation and mapping of a human gene (RPD3L1) that is homologous to
RT RPD3, a transcription factor in Saccharomyces cerevisiae." ;
RL Cytogenet. Cell Genet. 73:130-133(1996).
CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
CC -1- SUBUNIT: FORMS A COMPLEX WITH RBPAP4 AND ALSO WITH THE HISTONE
CC ACETYLTRANSFERASE PCAF AND THE ADAPTOR PROTEIN P300. ALSO FORMS
CC A COMPLEX WITH SIN3 AND SAP18. INTERACTS WITH TGIF.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC TISSUE SPECIFICITY: UBIQUITOUS, WITH HIGHER LEVELS IN HEART,
CC PANCREAS AND TESTIS, AND LOWER LEVELS IN KIDNEY AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; D50079; AAC50475.1; -
DR EMBL; D50403; BAA08909.1; -
DR MIM; 601241; -
DR InterPro: IPR000286; Hist_deacetylase.
DR Pfam: PF008050; Hist_deacetyl; 1.
DR PRINTS: PRO1270; HDASUPER.
DR PRINTS: PRO1271; HSDACETLASE.
KW Hydrolase; Nuclear protein.
FT CONFLICT 312 312 W -> R (IN REF. 2).
SQ SEQUENCE 482 AA; 55103 MW; 4B5587C1ED73B8D5 CRC64;

```

Query Match	56.7%	Score 1544	DB 1	Length 482
Best Local Similarity	59.4%	Pred. No. 2.6e+112		
Matches 291	Conservative 67	Mismatches 98	Indels 34	Gaps 7

[illegible]

```

OY      194 AAYADAWRTVSEFHKFGDVFPGCTGIIDQIGSGCYYSANPLDDGIDEESHLLFKPIM    253
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       186 AYTTTDRAWMTVSFHKXTCGEFFPTGTGLRDIOAGCKGYAVANNPLRGGIDDESEAIFRKVM    245
OY      254 GKVEIIRFGAVALQGCGADSISGRDLGCENLSIKHACVCYFMRSFNVLPLLGGCYYTI   313
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       246 SKVMEFPQSAAVVLCGSSDSISGRDLGCENLTKGHACVCYEFVKSFNLMILMGCCGYTII   305
OY      314 RNVAACMCYEITGAIVGEVEDMKPMHEHYETRYTGPDYTTLHAVPSNNENNKSQMLEIND   373
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       306 RNVACHMYETAIAVALDTLPINELPYNDIEFYEGPPDKHLSPSNMNTONTMLETEKIQR     365
OY      374 LLAHNLSKLQAHASVPFEGERPDTETPEV-DEDQEDGDKRMD-----PDSGM        418
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       366 LFENLRMLPHACGVCMQAIPDAIPAEESEDDEBDPOKRISTCSSDKTIALCEEESDSBE   425
OY      419 DVDDDRKKPIPS-----RVKREAVEPDTKDOKGLKGIMERGCCVEVDSEGSTKYTVGNP   473
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       426 FEGGCRKNNSNKPKAKRWTFE----DEKEKD-----PEEK---EVTEEERTKEKPRA    472
OY      474 VGVEEAASYKM 483
         ||| | | | | | | | | | | | | | | :||| | | | | | | | | | | | | | |
DB       473 KGVKE-EVNL 481

RESULT          5
HDAL_CHICK             STANDARD:           PRT:       480 AA.
AC      PS6517;
DT       15-JUL-1998 (Rel. 36, Created)
DT       15-JUL-1998 (Rel. 36, Last sequence update)
DT       20-AUG-2001 (Rel. 40, Last annotation update)
DE      HISTONE DEACETYLYASE 1 (HD1).
GN      HDAC1 OR HDAC1A
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX      NCBI_TaxId=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RL      Takami Y.;
RM      Submitted (DEC-1997) to the ENBL/genbank/DDBJ databases.
RR      [2]
RA      Sun J.M., Chen H.Y., Davie J.R.;
RL      Submitted (Feb-1998) to the ENBL/genbank/DDBJ databases.
RC      -|- FUNCTION: RESPONSIBLE FOR THE DEACETYLYATION OF LYSINE RESIDUES ON
CC      THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC      HISTONE DEACETYLYATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC      REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC      SIMILARITY).
CC      -|- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY).
CC      -|- SUPERCILIARY LOCATION: NUCLEAR (POTENTIAL).
CC      -|- SIMILARITY: BELONGS TO THE HISTONE DEACETYLYASE / ACDC / APPHA
CC      FAMILY. HD SUBFAMILY 1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the ENBL outposton --
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF039751; AAC96923.1; -.
DR      EMBL; AF043328; AAB99850.1; -.
DR      EMBL; AF044169; AAC00504.1; -.
DR      InterPro: IPR000286; His_deacetylase.
DR      Pfam: PF00850; Hist_deactyl_1.
DR      PRINTS; PR01270; HDASUPER.
DR      PRINTS; PR01271; HISDACETYLASE.
```

KM	Hydrolase; Nuclear protein.	POLY-Gly.
FT	DOMAIN	299 302
FT	DOMAIN	366 399
FT	DOMAIN	446 453
FT	CONFLICT	3 4
SO	SEQUENCE	480 AA; 54938 MW; 82C78CE285C79D9 CRC64;

Query Match	56.5%	Score 1538;	DB 1;	Length 480;
Best Local Similarity	57.9%;	Pred. No. 7.4e-112;		
Matches 282;	Conservative 74;	Mismatches 107;	Indels 24;	Gaps

QY	14	GVRKVCVEYDPPEVGNYYGGHGMKPHRIIMTHALLAHYLLOHMVILKPPAPAREDLG	73
Db	6	GTRKRVCYYDGVDGYNYGGGHGMKPHRIIMTHNILLNLGLYRKMEIYPRKANAEEMI	65
QY	74	RHHADYVSFLRSTTPETQDDIQQLKRFVNGECPYFDLSEFCQTAGSGVGSVKLN	133
Db	66	KYHSIDYIKFRLIRPNMSEYSKOMRFVNGECXPFDLEFEFCOLASGSAVASAVKLN	125
QY	134	HGLCDIAINMAGSLHAAKKCEASGFVYNDIVLAILELKQHERVLVYDIDIHKGAVEP	193
Db	126	KQGDIDIANMAGSLHAAKKSSAEGCVYNDIVLAILELLKYHQRLVLDIDIHKGAVEP	185
QY	194	AFYATDRMYTSFKFGDFPGTGHIDDTIGYSGRTYISLVNPILDGTIDESTHLKRPIM	253
Db	186	AFYTDRMYTSFKHYGEYFPGTDLDIDLAGKKYVAWVPPLRDGIDESTYAIRKPVI	245
QY	254	GKWEIFEPGAVLVOCGADSLSGRLCFNLSIGHAECVKPMKSFVNPILLGGGGTYI	313
Db	246	SKVMTETGPSNAVVLCCGSDLSIGRLCGFNLTIGKAHKCYEVKSPFLPYMLMGGGGYTI	305
QY	314	RNVARCWCYETGVALGYVEDMKPHEHYEYFPGDYTLHWAPSNMENKSNRMLEELRMD	373
Db	306	RNVARCWTYETAVALDTEIPNELPIYNDYFEYFPGDEFKLIHSPSNMTQMNTNEYLEIKOR	365
QY	374	LHNLSKTQHAPSPFPQPERPD-TETPEVDDEIDEDGDKRMD-POSIDVDVDDOKKPIPSRY	431
Db	366	LPENIRMLPHAPGVOMQIPDAVQEDSGDEEDEPKRSISRSDRRISCQ-----	417
QY	432	KREAVEPDTKCKDKLGIMRGKGCYEVEVDGSXTVTGVNPGVSEASYKMEEGNTKG	491
Db	418	EEFSDSEDGEDGGKRVANFKAKRAVKTEBE-----KEEEKKDEEKEXA	463
QY	492	GAEQAFF 498	
Db	464	KEEAEP 470	

RESULT	6
HD12_XENLA	STANDARD; PRT; 480 AA.
ID	HD12_XENLA
AC	042227;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	PROBABLE HISTONE DEACETYLASE 1-2 (HD1) (RBD3 HOMOLOG).
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxId=8355;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Patterson D., Wolfe A.P.;
RL	Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
CC	- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC	THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC	HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC	REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC	SIMILARITY).
CC	- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC	- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA

Query Match	55.9%;	Score 1521;	DB 1;	Length 480;
Best Local Similarity	59.1%;	Pred. NO. 1.6e-110;		
Matches 2/5; Conservative	77;	Mismatches 95;	Indels 18;	Gaps 4;

SEQUENCE FROM N.A.

SEQUENCE 400 MM; 24 / 4 / MM; 1B031022223DADB3 CRC04

[illegible]

```

Db      426  ECEGGRKMWANFKYKRVKTE-EKEEGDKMDVKREEKANEKTD 469

RESULT# 8
ID      HDAC2_MOUSE          STANDARD:          PRT:    488 AA.
AC      P70288:
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      HISTONE DEACETYLASE 2 (HD2) (Y11 TRANSCRIPTION FACTOR BINDING
DE      PROTEIN).
GN      HDAC2 OR YY1BP.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      NCBI_Taxid:10090;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      TISSUE=Lymphoma;
RX      MEDLINE-9/075080; PubMed-8917507;
RA      Yang W.-M., Inouye G.J., Zeng Y., Bearrs D., Seto E.;
RT      "Transcriptional repression by YY1 is mediated by interaction with a
RT      mammalian homolog of the yeast global regulator RPD3."
RL      Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC      -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC      THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC      HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC      REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC      SIMILARITY).
CC      -1- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC      AT LEAST WITH THE ZINC-FINGER TRANSCRIPTION FACTOR YY1.
CC      -1- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX WITH YY1.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR.
CC      -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC      FAMILY. MD SUBFAMILY 1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See license@isb-sib.ch).
CC      -----
CC      EMBL; U31758; AAC52889.1; .
CC      DR      MGD; MGI:1097691; Hdac2.
CC      DR      InterPro: IPR000286; His_deacetylase.
CC      DR      Pfam: PF008050; Hist_deacetyl; 1.
CC      DR      PRINTS: PR01270; HDASUPER.
CC      DR      PRINTS: PR01271; HISDACETLASE.
CC      KW      Hydroxylase; Nuclear protein.
CC      FT      DOMAIN 300 303      POLY-GLY.
CC      SEQUENCE 488 AA; 55302 MW; B9843D2A47A75157C CRC64;

Query Match      55.6%; Score 1513.5; DB 1; Length 488;
Best Local Similarity 55.7%; Pred. No. 6.1e-110;
Matches 280; Conservative 82; Mismatches 106; Indels 35; Gaps 7;

OY      8  LASGDGVKRRKVCYPPDEYGVNYYGGGHPMKPHRIIRMTHTALATYGLLOHNOVLKPPRA 67
DB      1  MAYSGGGRKKVCCYYDDDIGNYYGGGHPMKPHRIIRMTHTNLINYYGVRKKEIYRSPKA 60
OY      68  RERDLCRFHADYVSEFLKSIPTETQDDIIRLKRNVSEDCVDFGLSPFCOTYAGGSVG 127
DB      61  TAEETKTHSDIEYIFLKSIRPDNKKSEISKQORNVSEDCVDFGLSEFCOLSTGGSYVA 120
OY      128  GSVKTHNGLCDIAIMWAGGLHNAAKCEASGCVYNDIYLAILELLKQERVLVYVDIDINH 187
DB      121  GAVKLNROGTDMAVWAGGLHNAAKCEASGCVYNDIYLAILELLKYNORVLYIDIDINH 180
OY      188  GDGVAEAFVATDRVMTVSFHKFGDYFPGTGHIQDTGSGSKVYSLNVPDLDDSIDESYHL 247

```

```

Db 181 GDGVEAATYTTDRWTWTSFHKYGEFFPTGDLRLDGAGKGYAVNPFMRSGIDDESGQ 240
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 248 LFKPIIKVMYELFRGAVVLVLCGADSLSGDLRGCFNLSTIKHAECVCKPMRSFNNPLLLG 307
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 LFKPIIKVMYEMQPSAVVLVLCGADSLSGDLRGCFNLTVKHAKEVAKTFNNPLLLMG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 GCGTTIRNVAQCWCYETVALGVEYEDKMPHEHYEYFGPYTLVAVPSNNKNSRML 367
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 GGGYTIIRNVAQCWYETVALDCELPNELPYNDYFVFGPDPKHLISPSNNTNNTPEYM 360
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 368 EIRINDLHNLSKLOHAPSVFPFQEPPTETPEVED--VEEDGDKRMVDPDSMDVDDPK 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 EIKIQLRLEFENLRMLPHAFGVOMALPEPA----VHEDSGDEG---DPDK-----RI 406
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 426 PIPSKVKEAVEPPTKQD----GLKCMERKGE---VEVDESGTKVYGVNPGVC 477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 SIRASDKRIACDEFSDSEDEGGRRNVAADHKGAKKARIIEDEKKE-----E 455
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 EASVAMEEETNGKGAEGAFPPK 500
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 456 DKTDVKEEDKSKDNGSEKTDPK 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
HDAL_STRPU STANDARD; PRT: 576 AA.
AC P56518;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 1 (HD1).
GN HDAC1.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
CX NCBI_TaxID=7668;
RN [1]
RP Nemer M.;
RA SEQUENCE FROM N.A.
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch).
CC -----
Cc EMBL: AF032919; AAB87685.1; -
Cc InterPro: IPR000286; His_deacetylase.
Cc Pfam: PF00850; Hist_deacetyl_1.
Cc PRINTS: PRO1270; HDASUPER.
Cc PRINTS: PRO1271; HISDACETLASE.
Cc HydroLase: Nuclear protein.
Cc KW SEQUENCE 576 AA; 64078 MW; B3D11A844A2088E9 CRC64;

Query Match 55.4%; Score 1507; DB 1; Length 576;
Best Local Similarity 56.0%; Pred. No. 2,4e-109;
Matches 280; Conservative 75; Mismatches 119; Indels 26; Gaps 6;
14 GVKRRVCYFDPEVGNVYGGCHPMKPRIRIMTTHALLAHYGLLOHMVQLKDPFAERDLC 73

```

```

Db 5 GTRKRCVYDGDGVNYYGQGHPRMKPHRIKMTNHLNLVGLYRKMEIYRPHKAVMEBMT 64
QY 74 RFNADTVSLRSTITPETQODIROLKRFVNGDCVDFGLSYFCOTYAGSGVSKLN 133
Db 65 KYHSDYVFKLTIRPDNMSEYTKOMRFNVGDCVDFGLSYFCOTYAGSGVSKLN 124
QY 134 HGLCDIAINMAGLHNAKKEASGFCVNDIVATILELLKOHRYVYVDIDHGGVVE 193
Db 125 KOOTDIAINMAGLHNAKKEASGFCVNDIVATILELLKOHRYVYVDIDHGGVVE 184
QY 194 AFYATRVMTVSEHKFGDYFPGTGHIDIGYSGKYYSLNPLDDGIDDESHTLKPTM 253
Db 185 AFYTRVMTVSEHKFGDYFPGTGHIDIGYSGKYYSLNPLDDGIDDESHTLKPTM 244
QY 254 GKMEIFRPAVYVLOCADSLSDRLGCFNLSTKGHAECVKKFMRSNVPLLLGGGYIT 313
Db 245 CKVMEYVOPSAICIGCGASLSDRLGCFNLSTKGHAECVKKFMRSNVPLLLGGGYIT 304
QY 314 RNARCVCYETGVALGVEYEDKMPHEHYEYFGPDYTLHAPSNMKNKSRQMLEIRND 373
Db 305 RNARCVCYETGVALGVEYEDKMPHEHYEYFGPDYTLHAPSNMKNKSRQMLEIRND 364
QY 374 LHLNLSKLOHAPSVPRQEPPTETPEVDEDED-----GDKRMDPSDMV 420
Db 365 LYENMNMIRPHAPGVOMQPIPEDA-IPD-DSADAEANPDKRISIMADKRIGRDEFSO 422
QY 421 DDDRKPSPSVKREAVEDPTKDKGLKGMERKGCVEYEDSGSKYVCNPFVGEAS 480
Db 423 SEDEG--ETPLRPGGRHDSHRKAKRSKIDSPGKEADKSAKSPKAAK-AAAPOA 479
QY 481 VKMEEGTNKGAQAQFPK 500
Db 480 VPMD-----TTPAPPK 491

RESULT 10
HDAC2_HUMAN STANDARD: PRT: 488 AA.
AC 092759;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 2 (HD2).
GN HDAC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast;
RA MEDLINE=97075080; PubMed=8917507;
RT Yang W.-M., Inouye C.J., Zeng Y., Bears D., Seto E.;
RT "Transcriptional repression by YY1 is mediated by Rpr3."
RT mammalian homolog of the yeast global regulator Rpr3."
RL Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -1- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC WITH MAD, SIN3, YY1 AND N-COR.
CC -1- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; LOWER LEVEL IN BRAIN AND
CC LUNG.
CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31814; AAC50814.1;
DR MIM: 605164;
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl_1.
DR PRINTS: PR01270; HDASUPER.
DR PRINTS: PR01271; HISDACETYLASE.
KW Hydrolase; Nuclear protein.
FT DOMAIN 300
FT SEQUENCE 488 AA; 55325 MW; 3AE554ADC7734B70 CRC64;

Query Match 54.9%; Score 1494; DB 1; Length 488;
Best local similarity 54.3%; Pred. No. 2e-108;
Matches 277; Conservative 80; Mismatches 113; Indels 40; Gaps 6;

QY 8 LASGPDGVKRCVCFYDPENVNYYGQGHPRMKPHRIKMTNHLNLVGLYRKMEIYRPHKA 67
Db 1 MAVSQGGGKKKKYCYDDGIGNYYGQGHPRMKPHRIKMTNHLNLVGLYRKMEIYRPHKA 60
QY 68 RENDLCRFNADTVSLRSTITPETQODIROLKRFVNGDCVDFGLSYFCOTYAGSGV 127
Db 61 TAEEMKYHSDYVFKLTIRPDNMSEYTKOMRFNVGDCVDFGLSYFCOTYAGSGV 120
QY 128 GSVKLNHGLCDIAINMAGLHNAKKEASGFCVNDIVATILELLKOHRYVYVDIDH 187
Db 121 GAVKLNROOTMAVNMAGLHNAKKEASGFCVNDIVATILELLKOHRYVYVDIDH 180
QY 188 GGVVEAFYATRVMTVSEHKFGDYFPGTGHIDIGYSGKYYSLNPLDDGIDDESHTL 247
Db 181 GGVVEAFYATRVMTVSEHKFGDYFPGTGHIDIGYSGKYYSLNPLDDGIDDESHTL 240
QY 248 LRPPIKGMETFRPAVYVLOCADSLSDRLGCFNLSTKGHAECVKKFMRSNVPLLLG 307
Db 241 LRPPIKGMETFRPAVYVLOCADSLSDRLGCFNLSTKGHAECVKKFMRSNVPLLLG 300
QY 308 GGGYTRNARCVCYETGVALGVEYEDKMPHEHYEYFGPDYTLHAPSNMKNKSRQML 367
Db 301 GGGYTRNARCVCYETGVALGVEYEDKMPHEHYEYFGPDYTLHAPSNMKNKSRQML 360
QY 368 EETRNDLHNLKLOHAPSVPRQEPPTETPEVDEDEDKMD----- 413
Db 361 EETRNDLHNLKLOHAPSVPRQEPPTETPEVDEDEDKMD----- 410
QY 414 -PDSMDVDQDHRPIPSR--VKREAVEDPTKDKGLKGMERKGCVEYEDSGSKYVCN 469
Db 421 FDSDEDEGGGRNVDHKKAKARIEEDKETEKK-----TDVEEDSK-- 468
QY 470 GVNPFVGEASVYKMEEGTNKGAQAQFPK 499
Db 469 -----DMSGETDTRKTK-----SEQLSNP 488

RESULT 11
HDAC2_DROME STANDARD: PRT: 520 AA.
AC 094517; 017429;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE HISTONE DEACETYLASE (HD) (DRPD3).
GN RPD3 OR HDAC1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

```



```

OY 316 VARGCIEGVALGAEDEKMPHEHYEYEPFGDYTLHVAAPSMMENKNSROMLEIRNLL 375
    ||| ||| ||| ||| :||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 VSRCFTTETSAALAVELAINELPYNDYFFYFPGDFKLHISPSTMNTQNTSEYLEKIKNLP 364
OY 376 HNLKLIAPSPVFEGERPPTETPEVED-----QEDGDKRMDPSD-MDVDD-- 422
    || | || | || | || | || | || | || | || | || | || | || | || |
Db 365 ENLRLPLPAFCVOIALPEDAINDESDEDKVDKDDLRLQSODKKRIIPENEYSSEDEG 424
OY 423 -----DRKPISRVKREAVEPDTRKXDC-----LKGMCKCEVEYDESSGT 466
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 425 EGRGRDNRSYVGQRKRPRLDKTNSNKASSETTSSEIKDEKEGDADGESTANT 480

RESULT 12
HDAA2-CHICK STANDARD: PRT: 488 AA.
AC P56519;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 2 (HD2).
GN HDAC2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Takami Y.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC -I- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -I- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/ISB-
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF039752; AAC86924.1; -.
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl_1.
DR PRINTS: PR01270; HDASUPER.
DR PRINTS: PR01271; HISDACETYLASE.
DR KMW Hydrolyase; Nuclear protein.
DR DOMAIN 300 303 POLY-GLY.
SQ SEQUENCE 488 AA; 55153 MW; 4F79B9CD04A2D065 CRC64;
```


QY 188 GDVEEAFATDRTVTSFKFGDFFPCTGHIQDIGSGKYSINVLPLDGDIDDESYHL 247
 181 GDVEEAFATDRTVTSFKFGDFFPCTGHIQDIGSGKYSINVLPLDGDIDDESYHL 240
 QY 248 LKPRIMKWEIFERRAVNVLQCGADSLSGDRGCFNLISIKHAECYKFMSPVPLLLG 307
 241 LKPRIMKWEIFERRAVNVLQCGADSLSGDRGCFNLISIKHAECYKFMSPVPLLLG 300
 QY 308 GGGTTINVARCWCYETGVALGVEDEKMPHEHYEYFGPDYTLHAPSMWKNKSNQML 367
 301 GGGTTINVARCWCYETGVALGVEDEKMPHEHYEYFGPDYTLHAPSMWKNKSNQML 360
 QY 368 ELIRNDLHLNLSKLOHAPSVFPOERPPTETPEYDE--QEDGDRMPDSDMDVDDRK 425
 361 ELIRNDLHLNLSKLOHAPSVFPOERPPTETPEYDE--QEDGDRMPDSDMDVDDRK 406
 QY 426 PIRSRVREAVEPDTKDKD---GLKIMEKCKGE--VEYDESGSTVGTGVNPGVE 477
 407 PIRSRVREAVEPDTKDKD---GLKIMEKCKGE--VEYDESGSTVGTGVNPGVE 457
 QY 478 EASYKMEEGCNKGG 492
 458 KADYKEEDKSKDNMG 472

RESULT 13

HDA1_CAEEL

ID HDA1_CAEEL STANDARD; PRT; 461 AA.

AC 017695;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PROBABLE HISTONE DEACETYLASE 1.

GN HDA-1 OR C53A5.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodermidae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Mortimore B.;

RL Submitted (MOV-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON

CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).

CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL

CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA

CC FAMILY. HD SUBFAMILY 1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z81486; CAB03984.1; -

CC Mormprep; C53A5.3; CE08952.

CC InterPro: IPR000286; His_deacetylase.

CC Pfam: PF00850; Hist_deacetyl; 1.

CC PRINTS: PR01270; HDASUPER.

CC PRINTS: PR01271; HSDACETLASE.

CC K01271; HSDACETLASE.

CC HYDROLASE: Nuclear protein.

CC SEQUENCE 461 AA; 52137 MW; 9A104E88C5A1C07A CRC64;

Query Match 51.18; Score 1392; DB 1; Length 461;
 Best Local Similarity 57.24; Pred. No. 1, se-100;

Matches 259; Conservative 68; Mismatches 102; Indels 24; Gaps 4;

QY 1 MOTGGSNLSAGPCGVARKCYEFDPVGNATYYGQGPMPKPHRTIRMTALLAHGLQHQ 60
 1 MNSNGPLMEHG---KRAAYYDSNIGNYGQGVMPKPHRTIRMTALLAHGLQHQ 56
 QY 61 VLKPPAREEDLCRFADVDVSLRSTPTPTQDQJRLKRFVNGEDCVFPGLTSFCOT 120
 57 IIRFPFASFEEDMTFRFSDERTMTFLKSNANDNLKSPKQMLKRVNGDCPLFGLEYFCOL 116
 QY 121 YAGSVGYSYKLNHGLCDIAINNAGLHNAKCEASGFCYVDNYVALIELLKQERVLV 180
 117 SSGSLAATKTLKQKVYDIAINMGGILHNAKCEASGFCYVDNYVALIELLKQERVLV 176
 QY 181 VQIDIHGGDVEEAFATDRTVTSFKFGDFFPCTGHIQDIGSGKYSINVLPLDGI 240
 177 VQIDIHGGDVEEAFATDRTVTSFKFGDFFPCTGHIQDIGSGKYSINVLPLDGI 236
 QY 241 DDESYHLKPRIMKWEIFERRAVNVLQCGADSLSGDRGCFNLISIKHAECYKFMSPN 300
 237 TDVSYOSIFKPRIMKWEIFERRAVNVLQCGADSLSGDRGCFNLISIKHAECYKFMSPN 296
 QY 301 VPLLLGGGGYTLIRNVARCWCYETGVALGVEDEKMPHEHYEYFGPDYTLHAPSMNEN 360
 297 VPLLLGGGGYTLIRNVARCWCYETGVALGVEDEKMPHEHYEYFGPDYTLHAPSMNEN 356
 QY 361 KNSRQMLEIRNDLHLNLSKLOHAPSVFPOERPPTETPEYDE--DDEGDRMPDSDMDV 414
 357 ENSNDMLAKLQTVIANLELTPFVPSVOMPRIPEDLSALNDSLIAIXANPDKRLPDI 416
 QY 415 -----DSMDVDDRRKPIPSYKREA 435
 417 TDGMIDDDGGFTYGEREGGDRRN--ESDAKRAA 447

RESULT 14

HDA3_CHICK

ID HDA3_CHICK STANDARD; PRT; 428 AA.

AC P56520;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HISTONE DEACETYLASE 3 (HD3).

GN HDAC3.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC Takami Y.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON

CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).

CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL

CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA

CC FAMILY. HD SUBFAMILY 1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF039753; AAB96925.1; -

CC InterPro: IPR000286; His_deacetylase.

CC Pfam: PF00850; Hist_deacetyl; 1.

DR PRINTS: PRO1270; HOASUPER.
DR PRINTS: PRO1271; HISDACETLSE.
KW Hydroxylase: Nuclear protein.
SQ SEQUENCE 428 AA: 48901 MW: 42E32733AD2BBF07 CRC64

Query Match	50.68;	Score 1376;	DB 1;	Length 428;
Best Local Similarity	56.80;	Pred. No. 2.3e-99;		
Matches 242;	Conservative 74;	Mismatches 88;	Indels 22;	Gaps 4

QY	17	RKVCYFEPDPEVGNYYGCGHGMKPHRI:BMTHALLAHYGLQHMVOYLPPPARBERDLCRFH	76
Db	3	KTVAYFIDPDGNGFHYGAGHPMKPHRLALHSLVLAHTGLYKKMIVFKPYSQSDHMCRFH	62
QY	77	ADDYVSFLRSITPETOODI:RLQLRNFVNGEDCPVFDGLYSFCQTYAGSGWGSYKLMHGL	136
Db	63	SEDIYIDFLQGRVSPNNMOGFTKSLMAFNVDGDCPVFPGFLERCSHYTGASLQGAIVQLNKKI	122
QY	137	CDIAINNAAGCGLHAHKCKCBASGFCYVNDIVLAIEELKQHEVLYVDIDIHGGGVEEAFY	196
Db	123	CDIAINNAAGCGLHAHKCKCBASGFCYVNDIVLIELELRYHPRVLYLIDIDIHGGDVCFAFY	182
QY	197	ATDAWTVYSFHKFGDY--PPGTGHIDIDIGYGGKYSLNVPIDDCIDDESYHLFKPIMGK	255
Db	183	LIDRWTVYSFHKYGNFYFPFGTGMDEYVGAESGRYALNVP:LRDSIDDOQSYHLLFQPINQ	242
QY	256	VMEIFRPGAAVVLQCGADSLSGDRLGCTNLSTIKGAEVCKFRSEVNP:LLLGSGGYTRN	315
Db	243	VVDYTOPFCIVLQCGADSLGDRGRGCFMLSTRIGGECEYVKSFNIP:LLVLYGSGGYTRN	302
QY	316	VARQCVETGALGVENYEDMKRBEHEYEYFCPPDTLHY--APSNNEKNNSROMLEIENDL	374
Db	303	VAKMTYETSLVYDPAISSELPYSEYFEYFAPDFTLHPDVSTRILEONSRQYLDIOIQTI	362
QY	375	LHNLSKLOHAPSVEFOERP-----PDTE-----TPVEDQDQEDKRWDP	414
Db	363	FENIKMLNHASVQIHDVPSDLLSYDRDTPEDPBERGSEENYSRPEAANEFYDGDHDNDK	422
QY	415	DSQMDV 420	
Db	423	ESDVEI 428	

```

RESULT 15
HDA3_HUMAN          STANDARD;          PRT;          428 AA.
ID   HDA3_HUMAN
AC   O15379; O43268; Q9UEVO; Q9UE15;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 3 (HD3) (RPD3-2).
GN   HDAC3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE-T-cell, and Spleen;
RX   MEDLINE=98125547; Pubmed=9464271;
RA   Dangond F., Hafner D.A., Tong J.K., Randall J., Kojima R., Utku N.,
RA   Gullans S.R.;
RT   "Differential display cloning of a novel human histone deacetylase
RT   (HDAC3) cDNA from PHA-activated immune cells.";
RL   Blochem. Biophys. Res. Commun. 242:648-652(1998).
[2]
NP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE-Fibroblast;
RX   MEDLINE=98010646; Pubmed=9346952;
RA   Yang W.-M., Yao Y.-L., Sun J.-M., Davie J.R., Seto E.;
RT   "Isolation and characterization of cDNAs corresponding to an
RT   additional member of the human histone deacetylase gene family.";
RL   J. Biol. Chem. 272:28001-28007(1997).

```

RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE-98169450; PubMed-9501169;
 RA Emiliani S., Fischle W., Van Lint C., Al-Abed Y., Verdin E.;
 RT "Characterization of a human Rpd3 ortholog, HnAC3.";
 RN Proc. Natl. Acad. Sci. U.S.A. 95:2795-2800(1998).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE-99162406; PubMed-10051405;
 RA Makhnech U., Emiliani S., Najfeld V., Young S., Verdin E.;
 RT "Genomic organization and chromosomal localization of the human
 histone deacetylase 3 gene.";
 RN Genomics 56:197-202(1999).
 RN [5]
 RP SEQUENCE OF 95-353 AND 407-428 FROM N.A.
 RA Lynch E.D., Lee M.K., King M.-C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
 CC -1- FUNCTION: MAY PARTICIPATE TO THE REGULATION OF TRANSCRIPTION
 THROUGH ITS BINDING WITH THE ZINC-FINGER TRANSCRIPTION FACTOR YY1;
 CC INCREASES YY1 REPRESSION ACTIVITY.
 CC -1- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1/RPD3-2B (SHOWN HERE) AND
 CC 2/RPD3-2A: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / AFPA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U66914; AAC52038.1; -;
 DR EMBL: U75697; AAB88241.1; -;
 DR EMBL: U75696; AAB88240.1; -;
 DR EMBL: AF005482; AAB87752.1; -;
 DR EMBL: AF039703; AAC98927.1; -;
 DR EMBL: AF059650; AAC26509.1; -;
 DR EMBL: AF053138; AAC08351.1; -;
 DR EMBL: AF053137; AAC08351.1; JOINED.
 DR EMBL: AF053139; AAC08352.1; -;
 DR MIM: 605166; -;
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl1; 1.
 DR PRINTS: PRO1270; HDASUPER.
 DR PRINTS: PRO1271; HISDACETLASE.
 KW Hydrolase; Nuclear protein; Alternative splicing.
 FT VARSPIC 1 15 MAKVAVAYFDYDGVN -> MIVKPYGASQDMCK (1N
 FT ISOFORM 2).
 FT CONFLICT 359 359 R -> L (IN REF. 1).
 FT SEQUENCE 428 AA; 48847 MW; 94485C1EBDCF5A0D CRC64;
 Q0

Query Match	50.48;	Score 1373;	DB 1;	Length 428;
Best Local Similarity	56.68;	Pred. No. 4e-99;		
Matches 241;	Conservative 73;	Mismatches 90;	Indels 22;	Gaps 3;

QY	17	RKVCFFIDPEXGNTYYTQCGHMKPHRIKMHALLANTGLGHQVLRPFAPAREDLGRH	76
Db	3	KTAFVFPDPDGNHYHAGHMKPHRIALHSLVLAHGLKMKHIVFEPVQASOHMKRPH	62
QY	77	ADDVVSFLRSTPETQDDIROLKRFNVGDCVFPDLVYSFCQTAYAGSGVGSVKLNHGL	136
Db	63	SEDDIDLRVSPNNMGFTKSLNAFNVGDDCVFPELPEFGSHYTCAISLQGTQILNNKI	122

```

QY 137 CDIAINMAGGIHHAKKCEASGFCYVNDIVLALLELLKQHERVLYVDIDIHGSGVEEAFY 196
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 CDIAINMAGGLHHAKKFEASGFCYVNDIVLIGILELLKQHPRLVLYIDIDIHGSGVQEAFF 182
QY 197 ATRVMYVSFHAKGDI-FPGTGHIDIGYSGKYYSLNVPDLDDGIDDESYHLLFKPIMGK 255
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 LTRVMYVSFHAKYGNFFPGTGDMYEVGAESGRYYCLNVPDLRDIIDQSYKHLFOPVINO 242
QY 256 VMEIFRGAAYVLCGADSLSGDRLGCPNLSTKGHAECVKFMRSPNVPILLLGGGYTIRN 315
    ||:::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 VDPFQPTCTIVLCGADSLGCDRLGCPNLSTKRGCEVYKSFNIPLLVGGGGYTVRN 302
QY 316 VARGWCYETGVALGVEVEDKMPHEYYEYFPGDYTLHV-APSNNENKNSROMLEIRNDL 374
    |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 303 VARGWYETSLVEEATSEELPYSEYEFAPDFTLHPDYSTRIFNONSQYLDQIRQTI 362
QY 375 LHMUSKIQHAPSVYPOERPPDTET-----PEYDEQEDGDKRMDP 414
    |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 363 FENIKMLNHAPSVQIHGVPADLLTYDRIDEADAERKPREENYSRPEAPNEFYDGDHDNDK 422
QY 415 DSDMDV 420
    :||:::
Db 423 ESDVEI 428

```

Search completed: April 26, 2002, 17:57:50
 Job time: 333 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:57 ; Search time 25.88 seconds

(Without alignments)
1474.631 Million cell updates/sec

Title: US-09-645-337-2

Perfect score: 2722

Sequence: 1 MDTGNSLASGPDGVKRYKVC.....KMEEDCTNKGGAQAPPPKT 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: PIR_68:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2719	99.9	501	2 T05640	histone deacetylase
2	1986.5	73.0	513	2 T05640	probable histone d
3	1520	55.8	480	1 S60381	RPD3 protein homol
4	1392	51.1	461	2 T20163	hypothetical prote
5	1366	50.2	424	2 JC7102	histone deacetylase
6	1366	50.2	428	2 JC5834	histone deacetylase
7	1341	49.3	433	1 S22284	transcription regu
8	1336	49.1	405	2 T40300	histone deacetylase
9	1334	49.0	465	2 T23963	hypothetical prote
10	1298	47.7	419	2 T47443	hypothetical prote
11	1167.5	42.9	434	2 T11643	histone deacetylase
12	1077	39.6	507	2 T19067	hypothetical prote
13	1066.5	39.2	452	2 S64211	probable transcrip
14	475	17.5	375	2 D70388	acetylpolymyosin
15	471.5	17.3	470	2 S54089	hypothetical prote
16	448.5	16.5	389	2 E84054	acetylpolymyosin
17	443.5	16.3	387	2 S39643	acetylpolymyosin
18	425.5	15.6	367	2 T36278	acetylpolymyosin
19	367.5	13.5	158	2 T48929	hypothetical prote
20	352	12.9	359	1 B6266	acetylpolymyosin
21	320.5	11.8	706	2 S62933	hypothetical prote
22	317.5	11.7	380	2 D83174	probable acetylpol
23	301.5	11.1	425	2 G86217	protein T2767.14
24	299.5	11.0	142	2 T47441	hypothetical prote
25	297	10.9	310	1 A70481	acetylpolymyosin
26	296.5	10.9	1063	2 T46284	hypothetical prote
27	278	10.2	687	2 T43797	probable histone d
28	276.5	10.2	334	2 B75095	probable histone d
29	276	10.1	878	2 T17245	hypothetical prote

30	275.5	10.1	1095	2 T13964	probable histone d
31	271.5	10.0	338	2 H84173	acetylpolymyosin
32	264	9.7	335	1 H71071	hypothetical prote
33	261.5	9.6	310	2 B47050	glna 3'-region hyp
34	243.5	8.7	796	2 T32425	hypothetical prote
35	238	8.7	331	1 C69026	acetylpolymyosin
36	236	8.7	517	2 T27101	hypothetical prote
37	230.5	8.5	385	2 A81926	hypothetical prote
38	230	8.4	343	1 G64366	acetylpolymyosin
39	227	8.3	306	2 D82126	acetylpolymyosin
40	224.5	8.2	369	2 F81178	histone deacetylase
41	223.5	8.2	359	2 T05998	histone deacetylase
42	208	7.6	304	1 S74557	acetylpolymyosin
43	201	7.4	301	2 H75470	histone deacetylase
44	197	7.2	344	2 B83605	probable acetylpol
45	196.5	7.2	346	2 H83469	acetylpolymyosin

ALIGNMENTS

RESULT 1
T05640
histone deacetylase (EC 3.5.1.-) F20D10.250 - Arabidopsis thaliana
N:Alternate names: protein F20D10.250
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 12-Nov-1999
C:Accession: T05640
R:Bevan, M.; Medler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999
A:Reference number: 215420
A:Accession: T05640
A:Molecule type: DNA
A:Residues: 1-501 <BEV>
A:Cross-references: EMBL:AL035338; GSPDB:GN00062; ATSP:F20D10.250
A:Experimental source: cultivar Columbia; BAC clone F20D10
A:Genetics:
A:Gene: ATSP:F20D10.250
A:Map position: 4
A:Intons: 176/3; 323/3; 400/3; 424/2; 443/3; 467/3
C:Superfamily: RPD3 protein; RPD3/acuc homology
C:Keywords: hydrolase
F:21-320/Domain: RPD3/acuc homology <RAH>

Query Match	99.9%	Score 2719;	DB 2;	Length 501;
Best local similarity	99.8%	Pred. No. 1.1e-202;		
Matches 500;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MDTGNSLASGPDGVKRYKVCYFDPEYGNYYGOGHPMKRRIKTHALLAHGLLOHMO 60			
DB 1	MDTGNSLASGPDGVKRYKVCYFDPEYGNYYGOGHPMKRRIKTHALLAHGLLOHMO 60			
QY 61	VLPFPARERDLCRFHADDVVSFLRSTPTPTODQIRKRVNGEDCPFDLSFCOT 120			
DB 61	VLPFPARERDLCRFHADDVVSFLRSTPTPTODQIRKRVNGEDCPFDLSFCOT 120			
QY 121	YAGSGVGSVKLNHGLCDIAIMNAGGLHRAKCEASGFCVNDIVLLELLKQHERVLY 180			
DB 121	YAGSGVGSVKLNHGLCDIAIMNAGGLHRAKCEASGFCVNDIVLLELLKQHERVLY 180			
QY 181	VDDIHGDCVEAFYATDVMVVSFHKFGDYPPGCHIODIGSGKYYSLVNPIDDG 240			
DB 181	VDDIHGDCVEAFYATDVMVVSFHKFGDYPPGCHIODIGSGKYYSLVNPIDDG 240			
QY 241	DDESYHLLFPIMKVAEIRFPAVYLQCADSLSGDRLCFNLSTKGAECKPFRSFN 300			
DB 241	DDESYHLLFPIMKVAEIRFPAVYLQCADSLSGDRLCFNLSTKGAECKPFRSFN 300			
QY 301	VPLLILGGGTYTIRNVAKWCYETGVALGVEVBDKMEHEVYFGPDYTLHVAPSNMEN 360			
DB 301	VPLLILGGGTYTIRNVAKWCYETGVALGVEVBDKMEHEVYFGPDYTLHVAPSNMEN 360			

Query 361 KNSRQMLEEIRNDLLHNLISKLQHAPSVFQERPPDTPEYDEQDQDKRMDPSMDV 420
 |||||
 Db 361 KNSRQMLEEIRNDLLHNLISKLQHAPSVFQERPPDTPEYDEQDQDKRMDPSMDV 420
 |||||
 Query 421 DDDKRPISRKRAVEPDTKDKDLKIMRGKCEVEDESGSTVYGVNPGVEAS 480
 |||||
 Db 421 DDDKRPISRKRAVEPDTKDKDLKIMRGKCEVEDESGSTVYGVNPGVEAS 480
 |||||
 Query 481 VKMEEGTNKGAQCAQAPPKT 501
 |||||
 Db 481 VKMEEGTNKGAQCAQAPPKT 501
 |||||

RESULT 2
 T01413
 probable histone deacetylase (EC 3.5.1.-) rpd3 - maize
 C:Species: Zea mays (maize)
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 12-Nov-1999
 C:Accession: T01413
 R:Kossl, V.; Hartings, H.; Motto, M.
 Mol. Gen. Genet. 258, 288-296, 1998
 A:Title: Identification and characterization of an RPD3 homologue from maize (Zea mays L.)
 A:Reference number: 214321; MUID:98307342
 A:Accession: T01413
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-513 <ROS>
 A:Cross-references: EMBL:AF035815; NID:g265839; PIDN:AA050038.1; PID:g265840
 A:Experimental source: strain M22
 C:Genetics:
 A:Gene: rpd3
 A:Superfamily: RPD3 protein; RPD3/acuc homology
 C:Keywords: hydrolase
 F:27-325/Domain: RPD3/acuc homology <RAH1>

Query Match 73.08; Score 1986.5; DB 2; Length 513;
 Best Local Similarity 75.28; Pred. No. 5.7e-146;
 Matches 377; Conservative 44; Mismatches 61; Indels 19; Gaps 7;

Query 3 TGSNLSAS-GPDGKRYKCYFYDEPGYNYGGGCHPMKPHIRMTALLAHYGLLOHNOV 61
 |||||
 Db 8 SGGNSLPSVGPDGKRRKCYFYDVGNYVYGQGHMKPHIRMTALLAHYGLLOHNOV 67
 |||||
 Query 62 LKPPAREDLCKRHADYVSFLRSITPETQDDQIRQLKRNVEDCPVFDGLSFCQTY 121
 |||||
 Db 68 YRPNARERELCKRHAEEYINFLRSVTPETQDDQIRQLKRNVEDCPVFDGLSFCQTY 127
 |||||
 Query 122 AGGSVGVKLNHGLCDIAIMAGGLHAKKCEASGFCYVDIYALILELKHHERVLYV 181
 |||||
 Db 128 AGASVGVKLNHGLCDIAIMAGGLHAKKCEASGFCYVDIYALILELKHHERVLYV 186
 |||||
 Query 182 DIDIHGGVEEAFATRYMTVSFHKFGDYFPGTGHODIGYSGKYSYLNVPDLOGID 241
 |||||
 Db 187 DIDIHGGVEEAFATRYMTVSFHKFGDYFPGTGHODIGYSGKYSYLNVPDLOGID 246
 |||||
 Query 242 DESYHLKPLIMGKMYELFPGAVVLLCGADSLSGDLGCFNLSTKGAECKMRSEFNV 301
 |||||
 Db 247 DESYSLKPLIMGKMYEYFRGAVVLLCGADSLSGDLGCFNLSTKGAECKMRSEFNV 306
 |||||
 Query 302 PLLLLGGGCTTRNNARCMCYETGVALGVEDKMPREHEYYEYGPDTTLHVAAPSNMKN 361
 |||||
 Db 307 PLLLLGGGCTTRNNARCMCYETGVALGVEDKMPREHEYYEYGPDTTLHVAAPSNMKN 366
 |||||
 Query 362 NSRQMLEEIRNDLLHNLISKLQHAPSVFQERPPDTPEYDEQDQDKRMDPSMDV 421
 |||||
 Db 367 NSRQMLEEIRNDLLHNLISKLQHAPSVFQERPPDTPEYDEQDQDKRMDPSMDV 422
 |||||
 Query 422 DD-----RKPLPSRKRAVEPDTKDKDLKIMRGKCEVEDESGSTK---VTG 470
 |||||
 Db 423 DHKAVESSRSRSLIGIKIKREFGENATRYVDGGR-VASEHGLERPMADIGSSQAQAD 481
 |||||
 Query 471 VNPVGVEAS-VKMEEGTNK 490

Db 482 ASAMALDEPSNVNKNESSTK 502
 |||||
 RESULT 3
 S60381
 RPD3 protein homolog - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S60381; S43160
 R:Ladomery, M.R.; Lyons, S.; Sommerville, J.
 Submitted to the EMBL Data Library, December 1994
 A:Reference number: S60381
 A:Accession: S60381
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-480 <LAD>
 A:Cross-references: EMBL:X74844; NID:g773397; PIDN:CAA55211.1; PID:g602098
 C:Superfamily: RPD3 protein; RPD3/acuc homology
 F:13-312/Domain: RPD3/acuc homology <RAH1>

Query Match 55.88; Score 1520; DB 1; Length 480;
 Best Local Similarity 59.88; Pred. No. 6.8e-110;
 Matches 278; Conservative 73; Mismatches 96; Indels 18; Gaps 5;

Query 14 GVKRKYCYFYDEPGYNYGGGCHPMKPHIRMTALLAHYGLLOHNOVLPKPPAREDLCK 73
 |||||
 Db 6 GVKRKYCYFYDEPGYNYGGGCHPMKPHIRMTALLAHYGLLOHNOVLPKPPAREDLCK 73
 |||||
 Query 74 RPHADYVSFLRSITPETQDDQIRQLKRNVEDCPVFDGLSFCQTYAGGSVGVKLN 133
 |||||
 Db 66 KYHSDDYIKFLRSIRPNMSEYSKOMRFVNGEDCPVFDGLSFCQTYAGGSVGVKLN 125
 |||||
 Query 134 HGLCDIAIMAGGLHAKKCEASGFCYVDIYALILELKHHERVLYVDIDIHGGVEE 193
 |||||
 Db 126 KOOTDISVNMSSGLHAKKCEASGFCYVDIYALILELKHHERVLYVDIDIHGGVEE 185
 |||||
 Query 194 AFYATDVMVTSFHKFGDYFPGTGHODIGYSGKYSYLNVPDLOGIDDESYLEIFKPI 253
 |||||
 Db 186 AFYATDVMVTSFHKFGDYFPGTGHODIGYSGKYSYLNVPDLOGIDDESYLEIFKPI 245
 |||||
 Query 254 GKWEIIRPGAVVLLCGADSLSGDLGCFNLSTKGAECKMRSEFNVPLLLGGGCTT 313
 |||||
 Db 246 SKWEIIRPGAVVLLCGADSLSGDLGCFNLSTKGAECKMRSEFNVPLLLGGGCTT 305
 |||||
 Query 314 RNVARCMCYETGVALGVEDKMPREHEYYEYGPDTTLHVAAPSNMKNRQMLEEIRND 373
 |||||
 Db 306 RNVARCMCYETGVALGVEDKMPREHEYYEYGPDTTLHVAAPSNMKNRQMLEEIRND 365
 |||||
 Query 374 LHNLSKLQHAPSVFQERPPDT-----ETPEYDEQDQDKRMDPSMDV 423
 |||||
 Db 366 LFEURLPLPHAPQVQMAVADSIHDSGEDEDDPKRISRSSDKRIACDEFSSSD 425
 |||||
 Query 424 -----RKPLPS---RKRAVEPDTKDKDLKIMRGKCEVEYD 461
 |||||
 Db 426 EGEGGRKNVANFKKVKRYKTE-EKEGEDKKDKVKEEKAKDEKTD 469
 |||||

RESULT 4
 T20163
 hypothetical protein C53A5.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T20163
 R:Mortimore, B.
 Submitted to the EMBL Data Library, November 1996
 A:Reference number: 219232
 A:Accession: T20163
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-461 <ML>
 A:Cross-references: EMBL:Z81486; PIDN:CA03984.1; GSPDB:GN00023; CESP:C53A5.3


```

|||||
Db 193 LDRWMTVSFKHYGNYFFPGGDMYVAESGRYKYNLVPRLDGDSDSYKHLQPYVINO 242
256 VMEIFRPGAVYVLOCADSLSGDRLCFNLISIKHAECVKFMRSPVPLLLGGGGYITRN 315
243 VDFEYPCYVLOCADSLSGDRLCFNLISIKHAECVKFMRSPVPLLLGGGGYITRN 302
316 VARCWCTETGALVGEVDEKMPREHYEYFGPDYTLVH-APSNNMKNKSRMLEIRNDL 374
303 VARCWCTETSLVEALISEELPYSEYFAPDFTLHPDVSTRIENQNSRQYLDQILOTI 362
375 LHNLSKLOHAASVPRQERPPDTE-----PEVDEDEDGDKRMDP 414
363 FENLMLHAASVQIHDPADLLTYDRTDEADAERGPENYSREAPNEFYDGDHNDK 422
QY 415 DSDMDV 420
Db 423 ESDVEI 428

```

RESULT 7

```

S22284
transcription regulator RPD3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein N0305; protein YNL330C
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001
C:Accession: S22284; S51336; S51285; S55859; S57393; S63311
R:Vidal, M.; Gaber, R.F.
Mol. Cell. Biol. 11, 6317-6327, 1991
A:Title: RPD3 encodes a second factor required to achieve maximum positive and negative
A:Reference number: S22284; MUID:9204362
A:Accession: S22284
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-433 <VID>
A:Cross-references: GB:S66438; NID:q238961; PIDN:AB20328.1; PID:q238962
R:van Dyck, L.; Pascual-Ahuir, A.; Goffeau, A.
Submitted to the EMBL Data Library, December 1994
A:Description: A 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes
located gene for a putative aryl-alcohol dehydrogenase.
A:Reference number: S51334
A:Accession: S51336
A:Molecule type: DNA
A:Residues: 1-433 <VAN>
A:Cross-references: EMBL:X83226; NID:9642335; PIDN:CAA58228.1; PID:9642338
R:Nicaud, J.J.
Submitted to the EMBL Data Library, January 1995
A:Description: Sequence analysis of a 13.9 kb fragment of yeast chromosome XIV identifies th
A:Reference number: S51285
A:Accession: S51285
A:Molecule type: DNA
A:Residues: 1-201 <NIC>
A:Cross-references: EMBL:Z46259; NID:9633655; PIDN:CAA6368.1; PID:9854536
R:Matthi, M.; Nicaud, J.M.; Levesque, H.; Galliardin, C.
Yeast 11, 567-572, 1995
A:Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifies th
A:Reference number: S55859; MUID:95373280
A:Accession: S55859
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-201 <MAF>
A:Cross-references: EMBL:Z46259; NID:9633655; PIDN:CAA6368.1; PID:9854536
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R:van Dyck, L.; Pascual-Ahuir, A.; Purnelle, B.; Goffeau, A.
Yeast 11, 987-991, 1995
A:Title: An 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes as we
d gene for a putative aryl-alcohol dehydrogenase.
A:Reference number: S57391; MUID:96021610
A:Accession: S57391
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-433 <VAM>
A:Cross-references: EMBL:X83226; NID:9642335; PIDN:CAA58228.1; PID:9642338

```

```

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
R:Coster, F.; Jomiaux, J.L.; Goffeau, A.; Purnelle, B.; Van Dyck, L.
Submitted to the EMBL Data Library, April 1996
A:Reference number: S63151
A:Accession: S63151
A:Molecule type: DNA
A:Residues: 1-433 <COS>
A:Cross-references: EMBL:Z71606; NID:q1302450; PIDN:CAA96263.1; PID:q1302451; GSPDB:G
R:Matthi, M.; Nicaud, J.M.; Levesque, H.; Galliardin, C.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S63287
A:Accession: S63287
A:Molecule type: DNA
A:Residues: 1-201 <MAW>
A:Cross-references: EMBL:Z71606; GSPDB:GN00014; MIPS:YNL330C
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:RPD3; SD12; SD56; MIPS:YNL330C
A:Cross-references: SGD:S0005274; MIPS:YNL330C
A:Map position: 14L
C:Function:
A:Description: transcription regulation
C:Superfamily: RPD3 protein; RPD3/acuC homology
C:Keywords: nucleus; transcription regulation
F:23-322/Domain: RPD3/acuC homology <RAH1>

```

Query Match 49.3% Score 1341; DB 1; Length 433;

Best Local Similarity 55.6%; Pred. No. 4.3e-96;

Matches 230; Conservative 86; Mismatches 92; Indels 6; Gaps 2;

```

QY 16 KKKVCYFDEPVNYYGCGHPKPHIRMTTHALLAHYGLQHMVQLKPPAREDCRPF 75
11:|||||
Db 18 KRRVAYFYDADVGNVAVGAGHPKPHIRMAHNSILNMYGLYKMEYIRAPARPKQENCGF 77
11:|||||
QY 76 HADDYVFLNLTPTQDDIROLKRNVEDCPVFGGLYSCQYTAGSVGSVKLNHG 135
11:|||||
Db 78 HTDEYIDFLSVPTDNEMFKRESVKENVDGCPVFDGLYECISISGSGMEGAARLNG 137
11:|||||
QY 136 LCDIAIMAGGLHAKKCEASGFCYVNDIYALILELKHORVLYVYDIIHHDGVEAF 195
11:|||||
Db 138 KCDVAVNTAGGLHAKKSEASGFCYLVNDIVGLITELRYPRVLYTIDVHBDGVEAF 197
11:|||||
QY 196 YATDRWMTVSFKHYGNYFFPGGDMYVAESGRYKYNLVPRLDGDSDSYKHLQPYVINO 255
11:|||||
Db 198 YTTDRWMTVSFKHYGNYFFPGGDMYVAESGRYKYNLVPRLDGDSDSYKHLQPYVINO 257
11:|||||
QY 256 VMEIFRPGAVYVLOCADSLSGDRLCFNLISIKHAECVKFMRSPVPLLLGGGGYITRN 315
11:|||||
Db 258 IMEWYOPSAVVLQCGGDSLSDRLGCFNLISMEGHANCVNVKSPGIPMAYVGGGYITRN 317
11:|||||
QY 316 VARCWCTETGALVGEVDEKMPREHYEYFGPDYTLVHAPSNNMKNKSRMLEIRNDL 375
11:|||||
Db 318 VARCWCTETGLNNAVLDKDPINNEYTYGPDKLSVPSNNMKNKSRMLEIRNDL 377
11:|||||
QY 376 HNLKSLQHAASVPRQERPPDTE-----PEVDEDEDGDKRMDP 423
11:|||||
Db 378 ANLENTKVAASVQIHDPADLLTYDRTDEADAERGPENYSREAPNEFYDGDHNDK 431
11:|||||

```

RESULT 8

```

T40300
histone deacetylase (EC 3.5.1.-) - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: cryptic loci regulator protein clt6
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T40300; T43796
R:Lynch, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hillebert, H.; Mestl, D.; Duest
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21919
A:Accession: T40300
A:Status: preliminary; translated from GB/EMBL/DBJ

```



```

|||||:|||||: ||| ||: :|||: | |||
Db 319 LGGGYTLRNVARCMALETGYILGLRMDEIPGTSLSYHFTPLRLRNLPKKNDAWSA 378
OY 365 QMLEIRNDLHNLKQAHAPSVFQERPPDTEPEVNE-DOEGDKRMDDSDMDVDD 423
Db 379 AYLSIEKETIACLMKIGAPSVQNO---NIVGRIDEIQLIENENKSSKSSIEYE 434
OY 424 RKPIPSRYKREA-VEPDTKIDGLKIMERGKGEVEDESGSTKVTGVPVGEASVK 482
Db 435 VGKYSKMEECFCFVEEDSKPPSPFGQDPRRIQGYWGVDGRGLAPRSHSDV-IEEA--K 491
OY 483 MEESTNK 490
Db 492 YEDRRRK 499

RESULT 13
S64211
Probable transcription regulator YGL194C - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G1330
C:Species: Saccharomyces cerevisiae
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C:Accession: S64211; S62051
R:Bruschl, C.V.; Cogilevina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64183
A:Accession: S64211
A:Molecule type: DNA
A:Residues: 1-452 <BRU>
A:Cross-references: EMBL:Z72716; NID:g1322818; PIDN:CAA96906.1; PID:g1322819; MIPS:YGL19
R:Cogilevina, M.; Delneri, D.; Zaccaria, P.; Klima, R.; Bertani, I.; Bruschl, C.V.
submitted to the EMBL Data Library, September 1995
A:Description: A 6.7 kb fragment from chromosome VII of Saccharomyces cerevisiae contain
A:Reference number: S62051
A:Accession: S62051
A:Molecule type: DNA
A:Residues: 1-351; 'HSGTHS', 358; 'RIIHFI', 365-452 <COG>
A:Cross-references: EMBL:X91837; NID:g1177627; PIDN:CAA62950.1; PID:g1177634
A:Experimental source: strain FY1679
C:Genetics:
A:Gene: SGD:HOS2
A:Cross-references: SGD:S0003162; MIPS:YGL194C
A:Map position: 7L
C:Superfamily: RPD3 protein; RPD3/acuc homology
F:30-331/Domain: RPD3/acuc homology <RAH1>

Query Match 39.2%; Score 1066.5; DB 2; Length 452;
Best Local Similarity 48.9%; Pred. No. 8.1e-75;
Matches 204; Conservative 84; Mismatches 120; Indels 9; Gaps 5;

OY 18 KVCFFYDEVCNYYGCGHPKPRIRMTALLAHYGLQMOYKLPFAERDLCRPHA 77
Db 27 RVSHEFNKSVSHYHGVNHPKPRFLMTLDHLYSSYGLHKIMDLETYSARDELLQFHS 86
OY 78 DDVYSFLSIPPEQQDQIR-OLKRENYGECPPEDGILYSCQTGYAGGSVGSYKINGL 136
Db 87 EDYVNEFLSKVPENANKIPRGTELEFNIGDCCPIFONLIYDTTYTASLDAIRKLINQ 146
OY 137 CDIAINMAGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYDIDIHGDCVEAFY 196
Db 147 SDIAINMGGJHAAKKNPSGFCYVNDIVLAILNLRHPIRLVITIDIDLHGGDVGAEFY 206
OY 197 AIDRMVYSFHKF-GDYFPGTGHIDIGYSGKYSILNVPDLDDIDDSYLLFKPLMGK 255
Db 207 TTDVFTLSFKYNGEFPFGTDLTEICDCKGKHFALNVPLEDGIDDSYINLEKSIYDP 266
OY 256 VMEIRPCAVYVLOGADSLSDRLGCEMLSIKGAHECYKEMRSFNVPILLGGGGYTRN 315
Db 267 LITMKPLIIVQCGADSLGHRDLCCEFNLIKAGECYKFAKSCOLLVVGGGGYTRN 326
OY 316 VARCWCYETGVALGVEVDKMPH-EYVEFGPOVTLIVAPSNM-ENKNSROMLEIRND 373

```

```

|||||:|||||: ||| ||: :|||: | |||
Db 327 VSRLMVETETGLNDVLLPEDIIPDRDSFCGPDSYLYPMDDLVEKKNKLTLEDIR 386
OY 374 LHLNLSLOHAPSVFQERPPDTE-----TPVEDQEDGDKRMDDSDMDVDDRK 425
Db 387 CLENIRYLOAGAPSVRMDEACIPDIDISALTEBEDKLIQENMEETEAUSNREEMER 443

RESULT 14
D70388
acetoin utilization protein - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
C:Accession: D70388
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MID:98196666
A:Accession: D70388
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-375 <AOF>
A:Cross-references: GB:AE000719; NID:g2983517; PIDN:AAC07100.1; PID:g2983524; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: acuC1
C:Superfamily: acetylpolymine aminohydrolase; RPD3/acuc homology
F:6-306/Domain: RPD3/acuc homology <RAH1>

Query Match 17.5%; Score 475; DB 2; Length 375;
Best Local Similarity 32.4%; Pred. No. 3.9e-29;
Matches 116; Conservative 67; Mismatches 127; Indels 48; Gaps 10;

OY 26 EVGNVYGGCGHPKPRIRMTALLAHYGLQMOYKLPFAERDLCRPHADYVFLR 85
Db 11 DYGVYRPRKHPKPRIPVSLRLDLAMNLDDELKSPARKRELLFHTEDYINTLM 70
OY 86 STPEEQDQDQI---RLKRPNV-GEDCPVFDGLYSFCQTGYAGGSVGSYKLNGLCDIAI 141
Db 71 ---EAEKCCVPRGAREKNIGGYPENVSAMETGSLATGSTVQAIEEFLKG--NVAR 124
OY 142 NMAGGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYDIDIHGDCVEAFATDR 200
Db 125 NPAGGMHAKRSRANGFCYINDPAVGLIYLRKKGFKRLITLDLDRHCDGVQAFYDTQ 184
OY 201 VMTVSFHKFDY-FP-GTGHIDIGYSGKYSILNVPDLDDIDDSYLLFKPLMGKME 258
Db 185 VFLVSLQSEYAPRPFKFLIEIGBEKGGYNINIFLPGLMDNERLFLAEKSLIYKE 244
OY 259 IFRGAVYVLOGADSLSDRLGCEMLSIKGAHECYKMRSFNVPPL-----LLGGGGYT 312
Db 245 VFEDEVYVLDGTDPLLEDVLSKFNLS-----NVAFLLKAFINIVREYFGVYVIGGGYH 298
OY 313 IRNVARCMCYETGVALGVEVDKMPHEHYEYFGPOVTLIVAPSNMKNKSRMLEIR 370
Db 299 PYLARAKMTLWELCSGREVPEKL-----NNKAKELLSTI 333

RESULT 15
S54089
hypothetical protein YPR068c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YP9499.23c
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
C:Accession: S54089
R:Radcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54089
A:Accession: S54089
A:Molecule type: DNA
A:Residues: 1-470 <BAD>

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:52 ; Search time 21.38 Seconds

(Without alignments)
527.322 Million cell updates/sec

Title: US-09-645-337-2

Perfect score: 2722

Sequence: 1 MDTGNSLASGPDGVKRVKVC.....KMEEGTNKGAGAFPPPT 501

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2033.5	74.7	517	4	US-09-282-305-8
2	1950.5	71.7	439	4	US-09-282-305-6
3	1548	56.9	458	4	US-09-282-305-2
4	1530	56.2	482	1	US-08-528-255A-1
5	1530	56.2	482	1	US-08-717-365-1
6	1360	50.0	432	4	US-09-282-305-10
7	1130.5	41.5	351	4	US-09-282-305-4
8	253.5	9.3	335	4	US-09-446-504-80
9	106	3.9	754	4	US-09-214-564A-2
10	105	3.9	240	2	US-08-760-745-3
11	96	3.5	1435	2	US-08-568-459A-4
12	96	3.5	1435	2	US-08-487-826B-4
13	94.5	3.5	723	1	US-07-814-964-11
14	94.5	3.5	723	1	US-08-258-442-11
15	94.5	3.5	723	1	US-08-328-809-6
16	94.5	3.5	723	5	PCT-US92-11107-11
17	90.5	3.3	532	1	US-08-285-440-5
18	90.5	3.3	532	1	US-08-630-349-5
19	89.5	3.3	235	2	US-08-760-745-1
20	89.5	3.3	352	1	US-08-552-142A-4
21	89.5	3.3	352	1	PCT-US95-05741-4
22	89.5	3.3	352	5	PCT-US95-05741-4
23	88.5	3.3	674	1	US-08-803-973-7
24	88.5	3.3	674	1	US-08-803-972-7
25	88.5	3.3	707	1	US-08-803-973-12
26	88.5	3.3	707	1	US-08-803-972-12
27	88.5	3.3	1456	1	US-08-803-973-2

28	88.5	3.3	1456	1	US-08-803-972-2	Sequence 2, Appl
29	88.5	3.3	3052	2	US-08-557-122A-26	Sequence 26, Appl
30	88	3.2	262	1	US-08-106-981-6	Sequence 6, Appl
31	88	3.2	411	2	US-08-741-134-6	Sequence 6, Appl
32	87.5	3.2	258	4	US-08-961-083-90	Sequence 90, Appl
33	87.5	3.2	565	4	US-08-961-083-218	Sequence 218, App
34	87	3.2	1388	2	US-08-741-134-2	Sequence 2, Appl
35	87	3.2	1388	2	US-08-685-576-4	Sequence 4, Appl
36	86.5	3.2	237	2	US-08-760-745-5	Sequence 5, Appl
37	86.5	3.2	255	4	US-09-040-485-2	Sequence 2, Appl
38	86.5	3.2	764	1	US-08-375-300-4	Sequence 4, Appl
39	86.5	3.2	764	3	US-09-177-431-4	Sequence 4, Appl
40	86.5	3.2	764	5	PCT-US95-16930-4	Sequence 4, Appl
41	86.5	3.2	1089	1	US-08-375-300-2	Sequence 2, Appl
42	86.5	3.2	1089	3	US-09-177-431-2	Sequence 2, Appl
43	86.5	3.2	1089	5	PCT-US95-16930-2	Sequence 2, Appl
44	86	3.2	554	2	US-08-524-051-2	Sequence 2, Appl
45	86	3.2	554	3	US-09-052-778-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-282-305-8
; Sequence 8, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Zea mays
US-09-282-305-8
```

```
Query Match 74.7%; Score 2033.5; DB 4; Length 517;
Best local Similarity 76.8%; Pred. No. 2.2e-207;
Matches 385; Conservative 41; Mismatches 60; Indels 15; Gaps 6;

OY 3 TGGNSLAS-GPDGVKRVKVCYFDPEVGNYYGCGHPKPRIRIKHTALLAHYGLQHMOV 61
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 SGGNSLSPBPDQKRRVCYFDPEVGNYYGCGHPKPRIRIKHTALLAHYGLQHMOV 67

OY 62 LKPPAERPLRCRHADYVSFLRSIPETQODIROLKRFNVEDCPVPDGLYSCQTY 121
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 68 YRRPADRDLRCRHADYINFLRSVPTQODIROLKRFNVEDCPVPDGLYSCQTY 127

OY 122 AGGSVGSVKLNHGLCDIANNAGGLHAAKCEASGFCYVNDIVLALLLKHREVLVY 181
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 128 AGASVGAVALNHG-HDIAINMSGILHAAKCEASGFCYVNDIVLALLLKHREVLVY 186

OY 182 DIDHHHDVDEARVYALDRVNTVSFHKFGDYFRGTGHIODIGSGGYYSLNPLDDGID 241
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 187 DIDHHHDVDEARVYALDRVNTVSFHKFGDYFRGTGDIIRIGHSKGYSLNPLDDGID 246

OY 242 DESYHLLEKPIKMGVMEIFRPGAVVLOCGADSLSGDRLGCFNLSIKGHACVFMRSFN 301
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 247 DESYSLFLKPIKMGVMEIFRPGAVVLOCGADSLSGDRLGCFNLSIKGHACVFMRSFN 306

OY 302 PLLLLGGGGYTIIRNVAACWCYETGVALGVEDEKMPHEYYEYFGPDYTLHVAFPMNENK 361
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 307 PLLLLGGGGYTIIRNVAACWCYETGVALGVEDEKMPHEYYEYFGPDYTLHVAFPMNENK 366
```

OY 362 NSRQMLEIRNDLLHNLKLOHAPSVPPQRPPTETPEVDEQEDGDKRMDPSMDVD 421
 Db 367 NTRQLODDISKLNDLNLKLRHAPSVHFERVPTETPEQEDDQDDPDERHDDPSMDVD 426
 OY 422 DD-----RKRPRSVKREAVEPDTKDKGLKIMEKCKCEVEYDESGTK---VIG 470
 Db 427 DHKAVESRRSILGLIKREFGENATRYODGR-VASEHKGLEPMAEDIGSSKQAPQAD 485
 OY 471 VNPVGEAS-VKMEEGTK 490
 Db 486 ASAMADIDPSNVKNEPSSRK 506

RESULT 2

US-09-282-305-6
 ; Sequence 6, Application US/09282305
 ; Patent No. 6287843
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/282,305
 ; CURRENT FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 60/080,563
 ; PRIOR FILING DATE: 1998-04-03
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 439
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-282-305-6

Query Match 71.7%; Score 1950.5; DB 4; Length 439;
 Best Local Similarity 84.2%; Pred. No. 1.1e-198;
 Matches 358; Conservative 29; Mismatches 35; Indels 3; Gaps 3;

OY 4 GGNL-ASGPDYKRYCYFYDEPVGYGCGHPMKPHIRMTALHAYGLQHMVYL 62
 Db 8 GGNLPTTGGADGSKRRKCYFYDAEGNYYGCGHPMKPHIRMTALHAYGLQHMVYL 67
 OY 63 KPFARERDCRFHADYVSFLRSITPETQDQIRQLKRFNVEDCPVFDGLYSFCQTYA 122
 Db 68 RPHARERDCRFHADYVSFLRSITPETQDQIRALKRFNVEDCPVFDGLYSFCQTYA 127
 OY 123 GGSVGSVKLNHGLCDIAIMWAGGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYD 182
 Db 128 GGSVGSVAVKLNHG-HDIAIMWAGGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYD 186
 OY 183 IDIHGGVGEAFAATDRVTVSFHKGFDPGTHIODIGYSGKYYSLNVLDDGSDTD 242
 Db 187 IDIHGGVGEAFAATDRVTVSFHKGFDPGTHIODIGYSGKYYSLNVLDDGSDTD 246
 OY 243 ESYHLKPKPLMGKYMELFRPGAVVLOCGADSLSGDRLGCFNLSTKGAEVCYKMRSEFNP 302
 Db 247 ESYGSLKPKPLMGKYMELFRPGAVVLOCGADSLSGDRLGCFNLSTKGAEVCYKMRSEFNP 306
 OY 303 LLLGGGCVTIRNVARCWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAASNMEKN 362
 Db 307 LLLGGGCVTIRNVARCWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAASNMEKN 366
 OY 363 SRQMLEIRNDLLHNLKLOHAPSVPPQRPPTETPEVDEQEDGDKRMDPSMDVD 422
 Db 367 NTRQLODDISKLNDLNLKLRHAPSVHFERVPTETPEQEDDQDDPDERHDDPSMDVD 426
 OY 423 DRKP 427
 Db 426 DAKPL 430

RESULT 3

US-09-282-305-2
 ; Sequence 2, Application US/09282305
 ; Patent No. 6287843
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/282,305
 ; CURRENT FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 60/080,563
 ; PRIOR FILING DATE: 1998-04-03
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-282-305-2

Query Match 56.9%; Score 1548; DB 4; Length 458;
 Best Local Similarity 63.3%; Pred. No. 7.6e-156;
 Matches 280; Conservative 66; Mismatches 80; Indels 16; Gaps 3;

OY 7 SLASGPDYKRYCYFYDEPVGYGCGHPMKPHIRMTALHAYGLQHMVYLKPP 66
 Db 12 SPAGGEDAHRRRYSYFYDESIDYGGGHPMKPHIRHSHLVHYGLHRLLELSRPP 71
 OY 67 AREERDCRFHADYVSFLRSITPETQDQIRQLKRFNVEDCPVFDGLYSFCQTYAGSV 126
 Db 72 ASEADIRRFHSDYVAFNLASATGNPGVLDPRALKRFNVEDCPVFDGLYSFCQTYAGSV 131
 OY 127 GGSVVKLNHGLCDIAIMWAGGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYD 186
 Db 132 GAAYVKLNHGLCDIAIMWAGGLHAKKCEASGFCYVNDIVLAILLELKHREVLVYD 191
 OY 187 HGGVGEAFAATDRVTVSFHKGFDPGTHIODIGYSGKYYSLNVLDDGSDTD 246
 Db 192 HGGVGEAFAATDRVTVSFHKGFDPGTHIODIGYSGKYYSLNVLDDGSDTD 251
 OY 247 LKPKPLMGKYMELFRPGAVVLOCGADSLSGDRLGCFNLSTKGAEVCYKMRSEFNP 306
 Db 252 GLFQCIKKYMEYVYQPDYVLOCGADSLSGDRLGCFNLSTKGAEVCYKMRSEFNP 311
 OY 307 GCGGTTIRNVARCWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAASNMEKN 366
 Db 312 GCGGTTIRNVARCWCYETGVALGVEEDKMPHEHYEYFGPDYTLHVAASNMEKN 371
 OY 367 LEEIRNDLLHNLKLOHAPSVPPQRPPTETPEVDEQEDGDKRMDPSMDVD 426
 Db 372 LEEIRNDLLHNLKLOHAPSVPPQRPPTETPEVDEQEDGDKRMDPSMDVD 426
 OY 427 IPSRV-KREAVEPDTKDKGLK 447
 Db 417 QRSRLWGSAGAYSDTDDPSLK 438

RESULT 4

US-08-528-255A-1
 ; Sequence 1, Application US/08528255A
 ; Patent No. 5659016
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAMURA, YOSUKE
 ; APPLICANT: FURUKAWA, YOICHI
 ; TITLE OF INVENTION: RPDL PROTEIN AND DNA
 ; TITLE OF INVENTION: ENCODING THE SAME
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS, P.C.

```
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,255A
FILING DATE: September 14, 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terrylene F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 482
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung cDNA library
US-08-528-255A-1

Query Match          56.28; Score 1530; DB 1; Length 482;
Best Local Similarity 59.28; Pred. No. 6,8e-154;
Matches 290; Conservative 67; Mismatches 99; Indels 34; Gaps 7;

OY 14 GVKRKVCYFDPREYNGYGGGHPMKPHRTMTALLAHYGLQHMVLYKFPAREDDLC 73
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 GTRKRCYTYDGVGNTYGGGHPMKPHRTMTALLAHYGLQHMVLYKFPAREDDLC 65

74 RFHADYVSLRSTPTQODIQLKRFNVGDCPVFDGLYSCQTYAGSGVSGVKLN 133
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 KYHSDYIKFLRSTRPDMSEYKQMRNVGDCPVFDGLYSCQTYAGSGVSGVKLN 125

134 HGLCDIANNAGGLHAAKCCASGFCYVNDIVLAILLELKHQERYLYVDIDIHGDVEE 193
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 KQQTDIAVNMAGGJHAAKCCASGFCYVNDIVLAILLELKHQERYLYVDIDIHGDVEE 185

194 AFVATDRMYTSEFKPDYPRGTGTHIDIGSGKYKSLNPLDDGIDDESYHLLEFRPM 253
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 AFYTTDRMYTSEFKPDYPRGTGTHIDIGSGKYKSLNPLDDGIDDESYHLLEFRPM 245

254 GKVAEIRFPAVAVLQCGADSLSCDRLGCFNLISIKGHAECVFKMSFVNPULLLGGGVTI 313
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 SKVAKEMQPSAVVAVLQCGADSLSCDRLGCFNLISIKGHAECVFKMSFVNPULLLGGGVTI 305

OY 314 RNVAKCMCYETGVALGVEVEDKMPHEHYEYFDPDYTLHVAAPSMENKSKQMLEEIRND 373
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 RNVAKCMCYETGVALGVEVEDKMPHEHYEYFDPDYTLHVAAPSMENKSKQMLEEIRND 365

OY 374 LHLNLSKIQHAPSVYPPQERPDFTPEY-DEDDGDCGRND-----PPSDM 418
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 LFNLRMLPAPRGVQMOAIFPDALPEESGDEDDDPKRLISTGSSDKRIACEEESDSE 425

OY 419 DVDDDRKPIPS-----RVKRAVEPDTKDKDKGLGIMRGKGEVEYDESGSTKVTGVP 473
```

```
DB 426 EGEGRNNSNFKKARKVTE-----DEKEND-----PEEK-----EYTFEKKREKPEA 472
OY 474 VGVEASVKM 483
|:|||||:
DB 473 KGVE--EVKL 481

RESULT 5
US-08-717-365-1
; Sequence 1, Application US/08717365
; Patent No. 5763182
; GENERAL INFORMATION:
; APPLICANT: MAKAMURA, YUSUKE
; TITLE OF INVENTION: RPDL PROTEIN AND DNA
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOTTLE & TANIIS, P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,365
FILING DATE: 23-SEP-1996
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/528,255
FILING DATE: September 14, 1995
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terrylene F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 482
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung cDNA library
US-08-717-365-1

Query Match          56.28; Score 1530; DB 1; Length 482;
Best Local Similarity 59.28; Pred. No. 6,8e-154;
Matches 290; Conservative 67; Mismatches 99; Indels 34; Gaps 7;

OY 14 GVKRKVCYFDPREYNGYGGGHPMKPHRTMTALLAHYGLQHMVLYKFPAREDDLC 73
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 GTRKRCYTYDGVGNTYGGGHPMKPHRTMTALLAHYGLQHMVLYKFPAREDDLC 65

74 RFHADYVSLRSTPTQODIQLKRFNVGDCPVFDGLYSCQTYAGSGVSGVKLN 133
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 KYHSDYIKFLRSTRPDMSEYKQMRNVGDCPVFDGLYSCQTYAGSGVSGVKLN 125
```


Db 310 ORSLMSGAGVSDSDTDEPDSLK 331

RESULT 8

US-09-446-504-80
Sequence 80, Application US/09446504
Patent No. 6218150
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazuo
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/446,504
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 80
LENGTH: 335
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-446-504-80

Query Match 9.3%; Score 253.5; DB 4; Length 335;
Best Local Similarity 24.2%; Pred. No. 2,1e-18;
Matches 88; Conservative 71; Mismatches 155; Indels 49; Gaps 13;

QY 22 FYDEYVGNYYGGHGHKPKHRIKTHALLAHYGLDHQMOYLKPPRAERDLGRHADDYV 81
DB 6 FYSKKFFNLHRPSEYHNPENPKRLTIVLSKRELGLEERIE--EPNPVEETVEKIHHDYI 63
DB 82 SFLRSITPETQDQIRQLKRFNVEDCPVFDGLYSFCQYAGSGVSGSVK-LNHGLCDIA 140
DB 64 NFVK-----EAEVKITRL-----DPDTYVSGTWSAALLALGAARSAALAHGGLHMA 114
QY 141 INMAGLHNAKCEAS-----GFCYNDIVLAILLELKH-ERYLVYDIDIHNGDYEEA 194
DB 115 LVRRPGHHAGRGRCAMCAPTLGFCIFNNMAAVVTLKEGVGKVVVIDPDHNGNGTOEI 174
QY 195 FYADRWATYSFHKFGYFPGTGHIDIGSGKYYT-LNVPDLDDGIDDSYHLLEKPYIM 253
DB 175 FWNPDVYTHIDLH-RDIYPSGDSVEVG-GSNAYGSKINLPMRYSGDDYIYVWDEIV 232
QY 254 GKVMETIRPGAVVYOCGADSLSGDRLGCFNLSTIKGHAECVFMFSFNVPLILGGGGYTI 313
DB 233 LPIYEEKPKYIVISAFDFGKGGTLTLRLTFESFYACATLKKYPLAFLTEGGYVGL 292
QY 314 RN-VARCMCEYTGVALGVEEDMPREHYEYFPGDYTLHVPASNNMKNNSROMLEIR 371
DB 293 DKGPAPITMGYEGKA-----KAREPRYE-----TLKLA-----EYK 326
QY 372 ND 374
DB 327 DTL 329

RESULT 9
US-09-214-564A-2
Sequence 2, Application US/09214564A
Patent No. 6150515
GENERAL INFORMATION:

APPLICANT: Sharp, Phillip A.
APPLICANT: Zhou, Qiang
TITLE OF INVENTION: TAT-SF: Cofactor For Stimulation Of Transcriptional
Elongation By HIV-1 TAT
FILE REFERENCE: M0656/77042
CURRENT APPLICATION NUMBER: US/09/214,564A
CURRENT FILING DATE: 1999-08-18
PRIOR APPLICATION NUMBER: US 60/021,218
PRIOR FILING DATE: 1996-07-03
PRIOR APPLICATION NUMBER: US 60/033,152
PRIOR FILING DATE: 1996-12-13
PRIOR APPLICATION NUMBER: PCT/US97/11713
PRIOR FILING DATE: 1997-07-03
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 754
TYPE: PRT
ORGANISM: Homo sapiens
US-09-214-564A-2

Query Match 3.9%; Score 106; DB 4; Length 754;
Best Local Similarity 20.4%; Pred. No. 0.04;
Matches 80; Conservative 35; Mismatches 129; Indels 148; Gaps 15;

QY 230 YSLNVLDDGIDDESHTLLEKPIMGKMEIFPGAV-----VLQCGADSLSGDLGCF- 282
DB 136 YVGLPDPDVIDE-----FIQLMSKFGITMRDPQTEEFKVLKYNOCNLKGDGLCYL 189
QY 283 -----NLST-----KGHAECVFMFSFNVPLILGGGGYTIIRNARCWCETGY 326
DB 190 KRESVELATLLEDDELKGYKLHVEAKFO-----LKGEDASKKKKKCKDKYKKL 240
QY 327 ALGVEVEDKMPREHYEYFPGDYTLH---VAPSN---MENKSRMLEIRNDLHNL 379
DB 241 SMOQKQDLMRPER-----AGPSMRHERVYILKNMHPHOFEDDPLVNEIRHDLVEES 296
QY 380 KLGHPASVPRQERPPDTETREVEDEQEDG-----KRPDPSMDVD 421
DB 297 KFGQIRKLFLDHRHPGVASVSFRDPEADYCIQTLDGRWFGROITTAAMDGTTOYVE 356
QY 422 DD-----RKPIPSVYKRAVEVPD----- 440
DB 357 ETSREBERLGRDEAFLEAPNARGLSVQTLSLKRAGPSRAHSEHSTSMNAOETA 416
QY 441 -----KDKDLKGMERG-----KCEYEVDSGSTKYT 469
DB 417 TGMAFPEIDKKFKPKTFDG--GEFEGASENNNAKESSPKRAFEQCPKESEEGCPK-- 472
QY 470 GVNPDVGE-EASYKMEBEGTNKGADQAPPPK 500
DB 473 ---RGFEGSCSQKESSEGNPVGSEDSPPK 500

RESULT 10
US-08-760-745-3
Sequence 3, Application US/08760745
Patent No. 5972658
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Murty, Lynn E.
TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,745
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0169 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 598956
US-08-760-745-3

Query Match 3.9%; Score 105; DB 2; Length 240;
Best Local Similarity 23.8%; Pred. No. 0.0075;
Matches 48; Conservative 27; Mismatches 77; Indels 50; Gaps 7;

QY 334 DKME-----HEYYE-YGSPDTLHAPSNM-----ENKNSQMLEIR 371
DB 31 DEMEAAVKSTANKYQVFFGTHETAFLEPDLFPYSESEKKEKPKKKKFSGLWEIE 90
QY 372 NDLLHNLSTLOHAPSVFPOERPTTPEVDEDEGDGDKRMDPSDM-----VDDRK 425
DB 91 NNPVAKSGVSSQSKSCVEEP-----EPPEALAGDGDGKGNAGSSDEGKLVIDEPK 146
QY 426 -----PIPSRVKREAVDPDTKDKGLKGMERGKCEVEVDSEGSTKY 468
DB 147 EKNKGAALKRRAGDLLDSDPKRPF-EAENPEGEKEKATLEVERPLMEYEKNSPSEPG 205
QY 469 TGVNPGVEEASVKMEEGTKK 490
DB 206 SGRGPPQEEDEEBEATKE 227

RESULT 11
US-08-568-459A-4
Sequence 4, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: SIm, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knodde Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-568-459A-4

Query Match 3.5%; Score 96; DB 2; Length 1435;
Best Local Similarity 18.6%; Pred. No. 1.4;
Matches 76; Conservative 61; Mismatches 155; Indels 116; Gaps 19;

QY 169 LELLKQHERLYVDIDIHGDGEAEAYATD-----RVNTVSFHKGFPGTGHIDIG 223
DB 509 LMKTEH-----ILAIAYESRILAKRYKKNDKDEYCKIINKTFADRLDIDGTDYNDLS 564
QY 224 -----YSGKYYSLNPLDGDIDESTYHLLFKPIMGKMEIFPPGAVVLOCGADSL 275
DB 565 NRKLVGKINTNSKYVRHKKNDKLFDEWKKYIKKQVWVYISWVFDKTV---CKEDDIE 621
QY 276 GDLGCFNLSTKHAACVCKPMSFNPLLLGGGY-----TIR----- 314
DB 622 -----NIPQFRFSE-----WGDDYCGDKTKMIETLKYCEKPCEDDN 661
QY 315 NVARCWCYETGVALGVEDEKMEHEYYEY-FGPDYTLH-----VAPSMENKNSR---- 364
DB 662 CKSKCNYSKEMISKKEEYKQAK-QYQYQKQNNYKMYSEFKSIPEVYLKYSKCSN 720
QY 365 -----QMLEIRNDLLHNLSTLOHAPSVF-----QERPPD-----TETPEV 401
DB 721 LNFEDFKELHSDYKNTCTMCPKVDVPISILRNNEQTSQEAVPENTELIARHETTPSI 780
QY 402 DEQDEGDKRMDPSDMVDVDDRRKPIPSRVKREAVDPDTKDKGL-----KGI 449
DB 781 SEGPK-GNEOKERD-----DDSLSKISVSPENRPETDAKDTSNLKLKGVDISMRAV 834
QY 450 MERGKCEVEVDSEGSTKYGVN--PVGVEASVKMEEGTKKGAEO 495
DB 835 IGSSPDNINVTEDGD-NISGVNSKPLSDVDPRDKKLEEDONDSDEE 881

RESULT 12
US-08-487-826B-4
Sequence 4, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: SIm, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.

RESULT 15
US-08-328-809-6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:52 ; Search time 43 Seconds

(Without alignments)
863.040 Million cell updates/sec

Title: US-09-645-337-2

Perfect score: 2722
Sequence: 1 MDTGNSLASGPDGVKRVKVC.....KMEDEGTNKGAGQAPPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A.Geneseq-1101.*
2: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2722	100.0	501	22	AA67811
2	2722	100.0	501	22	AA67815
3	2138	78.5	473	21	AA158829
4	2033.5	74.7	517	20	AA288800
5	2005	73.7	493	21	AA588828
6	1950.5	71.7	439	20	AA28799
7	1568.5	57.6	471	21	AA605187
8	1568.5	57.6	471	21	AA639372
9	1560.5	57.3	471	22	AA67812
10	1548	56.9	458	20	AA28797
11	1544	56.7	482	18	AAW29324

12	1544	56.7	482	22	AA67164	Human histone deac
13	1544	56.7	482	22	AA67815	Human histone deac
14	1544	56.7	488	21	AA568985	Human prostate can
15	1530	56.2	482	17	AA688919	RPDL, a human foet
16	1494	54.9	488	22	AA649955	Ruman histone deac
17	1421.5	52.2	432	21	AA605188	Arabidopsis thalia
18	1421.5	52.2	432	21	AA639373	Arabidopsis thalia
19	1412	51.9	370	22	AAU00243	Human histone deac
20	1382.5	50.8	425	21	AA605189	Arabidopsis thalia
21	1382.5	50.8	425	21	AA639374	Arabidopsis thalia
22	1373	50.4	428	22	AA649956	Human histone deac
23	1360	50.0	432	20	AA628801	Maize histone deac
24	1341	49.3	433	22	AA67166	Yeast histone deac
25	1130.5	41.5	351	20	AAV28798	Maize histone deac
26	855.5	31.4	415	22	AAW40374	Human polypeptide
27	852	31.3	377	21	AAV83027	Human histone deac
28	852	31.3	377	22	AA684361	Amino acid sequenc
29	629	23.1	180	22	AA670496	Saccharomyces cere
30	566.5	20.8	296	22	AAW42160	Human polypeptide
31	494.5	16.2	120	21	AAV58830	Wheat histone deac
32	483	17.7	133	18	AAW29329	A partial deacetyl
33	477	17.5	375	22	AAU00241	Cys75Ser/Cys75Ser
34	475	17.5	375	22	AAU00242	Wld-type histone
35	471	17.3	375	22	AAU00242	Tyr297Pine mutant o
36	456	16.8	125	18	AAW29330	Arabidopsis thalia
37	333.5	12.3	398	21	AA607246	Arabidopsis thalia
38	328.5	12.1	468	21	AA651592	Arabidopsis thalia
39	328.5	12.1	499	21	AA651591	Arabidopsis thalia
40	328.5	12.1	552	21	AA651590	Arabidopsis thalia
41	327.5	12.0	364	21	AA607248	Arabidopsis thalia
42	327.5	12.0	377	21	AA607247	Arabidopsis thalia
43	321	11.8	69	18	AAW29326	Hpx protein family
44	320.5	11.8	468	21	AA608532	Arabidopsis thalia
45	320.5	11.8	499	21	AA609531	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ID AAB67811 standard: Protein; 501 AA.
XX
AC AAB67811;
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated ATRPD3A.
XX
KW Histone deacetylase; ATRPD3A; RPD3; gene expression; transgenic plant;
KW HDAL; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRI & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Trian L, Wu K;
DR MPI; 2001-258457/27.
XX
DR N-PSDB; AAF80350.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase

XX Claim 10; Fig 1A; 91pp; English.
PS

CC The present sequence represents *Arabidopsis thaliana* histone deacetylase
CC designated AtRpd3A. The protein is homologous to yeast RPD3 and HDAC1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operatives
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.

501 AA;
Sequence

Query Match	100,0%	Score 2722	DB 22	Length 501
-------------	--------	------------	-------	------------

Best Local Similarity 100.0%; Pred. No. 3.7e-275;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY	1	MDTGNLSIASCPDPCVKRVCYFYDPEVGNATYCGGHHMKPHRIEMTALLAHNGLLQHM	60
Dd	1	mdtgnslasagpdpvkrtkvcfydpevgnyyyqgghbmphritmthalahngllqhm	60
QY	61	VLAFFPARERDLCFPHADDVYSFLRSTTRPETQDQIQOLKRFVNGEDCPYFDGLYSFCOT	120
Dd	61	vlkfpfrrerdlcfrhaddvysflrstltpetqdqirqlkrfingedcprfdglysfct	120
QY	121	YAGSGVGSVYLNHGLCDIAIMNAGGLHHAKKCEASGFCVYNDIVLATELLKQHEHYEY	180
Dd	121	yagsgvgsvylnhgldialmnagglhhakkceasgfcyvndivlatellkqhehy	180
QY	181	VDIDIHNGDVEEAFYATDRVMYSPHKFGDYPFGTHIODIGTGSKRYSLNVPILDGT	240
Dd	181	vdidihngdveeafyatrvmysphkfgdyfpgthiodigsgskryslnvpilddgt	240
QY	241	DDSEYHLLEFKIMCKVMEIFRPGAVVLQCGADSLSGRLGCFNLSIKGHAECKFMKFSN	300
Dd	241	ddeeyhlhlfkrlmkvmeifrgpavvlqcgadslsgrlgcfnlslkghaeckfmrfsn	300
QY	301	VPLLLLGGCGGTTIRNVARCMWCYETGVALGVEEDKMEHEHYEYFGDYTLHVAPSNMEN	360
Dd	301	vpllllggggtylrnvarcwcyetgvalgveedkmehehyeyfgpdytlhvasnmn	360
QY	361	KNSKQMLKEEIRNDLIHLNLSKIOHNASVPPOERPRDTRTPVEDGQFNGDKRMPPDSMDV	420
Dd	361	knsqmlkeelrndlhlhlnlskiohnasvppocerprdrtrpvedeqfngdkrmppdsmdv	420
QY	421	DDDRKRLPSRYKRAVEPDRDKDKGLGIMERGKGEVEVDESGSTAVTVNPNVGYEAS	480
Dd	421	dddrkrlpsrykravepdrdkdkglgimeyrgkgevevdesgstktvgnpnvygeas	480
QY	481	VKMEEEGTSKNGAGAEAPPPKT 501	
Dd	481	vkmeeeegtskngaeapppkt 501	

RESULT 2
AAB67165

ID AAB67165 standard; Protein; 501 AA.

AC AAB67165;

DT 12-APR-2001 (first entry)

DE Arabidopsis histone deacetylase 1 HDAC1.

KW Gene expression; disease model; treatment; suppression;

KM chromatin inactivation; DNA binding site; histone deacetylation
 XX
 OS *Arabidopsis thaliana*.
 vv

PN WO200102019-A2.

PD 11-JAN-2001

PF 28-JUN-2000

PR 30-JUN-1999; 99GB-0015126.

PA (IMCO-) IMPERIAL COLLEGE INN

PI Buluwela L, Ali S;

... WPI; 2001-138069/14
DR

PT Suppressing express

to or associated with increasing peptide

PT -
XX
PS
XX
Disclosure: Fig 5: 65bp; English.
XX
CC The present invention describes a method of suppressing the expression of
CC a gene of interest by introducing into the cell a protein containing a
CC DNA binding site and a chromatin inactivation portion, or a nucleotide
CC encoding such a peptide. Preferably the chromatin inactivation portion is
CC part of the histone deacetylation (HDAC) complex. The method can be used
CC in disease treatment, for example in the treatment of cancer by the
CC suppression of oncogenes, and in the production of disease models.

5Q	Sequence	501 AA.
----	----------	---------

Query Match	Score	DB	Length
100.0%	2722	22	501

Best Local Similarity 100.0%; Pred. No. 3.1e-2/5;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY	1	MDTGNLSLSPGCVKRRVCYFUDREGVGNYYUGGCHPMKPRIIMFTAHLLAHNGLLQHMQ	60
Dp	1	mdtgnslasppgcvkrrvcyfudrevgnyyuggchpmkpririmtahllahngllqhmq	60
QY	61	VLKPFPAKERDLCRFHADDVYSLRSITPETODOLBOLKRFNVGEDCPVFDGLYSCQT	120
Dp	61	vlkpfparerdlcrfhaddvyslrsltpetodolqlkrfnvgedcpvfdglsytcql	120
QY	121	YAGSGVGSVYLNLNGLCDIALNMAAGLHHNAKCSASPCVYNQIVYALILELLKHQHEVLY	180
Dp	121	yagsvgsvylnlncgldialnmaaglhhnakcsaspcvynqivyalilellkqhenvly	180
QY	181	VDIDILHHQDVEEAPAYADRMYTSFHKFGDYFPGTGHIDIGYSGSKYYSLANPLRDLGI	240
Dp	181	vdidilhhqdveeapayadrmytsfhnkfgdyfpgtghidigysgskyysslanplrdldgi	240
QY	241	DDESEHLLFKRIHMGKMEIFRPGAVYLQCGADSLSGDRLCGFNLSTKGHAECYFKMFSFN	300
Dp	241	ddesehllfkrihmgkmeifrpqavvylqcgadslsgdrjgcfnlstkghaecvfkfmsfn	300
QY	301	VPLLLLGGCGGTIRNVARCWCYERFGVALGYVEERKMPREHNEYEYFGFDYTLHVAFSMEN	360
Dp	301	vpllllgggyltrnvarcwcyeerfgvalgyveerkmprehneyeyfpgdytllhvafrsmen	360
QY	361	KNSKQMLEELFNOLLNLHLSKIQHAPSVRQDERPRDTEPREYDEQDEGDKRMQDPSIMDY	420
Dp	361	knsqmlleelfnollnlhlskqhapyvrqderprdtetpreydeqdeqdkrtwdpsimdy	420
QY	421	DDDRKPIPSRVKREAVRPDTEKQKGLGIMRGKGCVEVEDESSTKVTGVCNRPVGEAS	480
Dp	421	dddrkpiipsrvkreavrpdtckqglgimrgkgcvevedesstkvtygcnrvpgeas	480
QY	481	VKMEEGTNGKGAEDAPPKPT 501	

Db 481 vkmeegtnksgaefpckt 501

RESULT 3

AAV58829 standard: Protein: 473 AA.

AAV58829; 08-MAY-2000 (first entry)

XX Soybean histone deacetylase 1 (HD1) protein.

XX Chromatin associated protein: histone deacetylase gene 1; HD1.

KW soybean; transgenic plant; transcription regulation.

OS Glycine max.

PN WO200004177-A1.

PD 27-JAN-2000.

PF 13-JUL-1999; 99WO-US15807.

PR 14-JUL-1998; 98US-0092841.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Cahoon RE, Vollmer SJ;

XX WPI: 2000-182439/16.

DR N-PSDB; AAZ58260.

PT New nucleic acid fragment useful as probes and primers, for transforming plants

PS Claim 1: Page 27-28; 36pp: English.

XX The present sequence is that of soybean histone deacetylase 1 (HD1), a chromatin associated protein, as deduced from a soybean root cDNA clone (see AAZ58260) isolated on the basis of homology to plant histone deacetylases. The invention relates to isolated rice, soybean and wheat nucleic acid fragments encoding HD1. It also relates to the construction of a chimeric gene encoding all or a portion of HD1, in sense or antisense orientation, where expression of the chimeric gene results in production of altered levels of HD1 in a transformed host cell. The availability of nucleic acid sequences encoding (portions) of histone deacetylase proteins will facilitate studies of global transcriptional regulation in eukaryotic cells, and will also provide mechanisms to control transcriptional gene regulation in plants.

SQ Sequence 473 AA;

Query Match 78.5%; Score 2138; DB 21; Length 473;

Best Local Similarity 80.7%; Pred. No. 3.5e-214; Matches 392; Conservative 33; Mismatches 37; Indels 24; Gaps 2;

QY 1 MDTGNSLACPDGVRKRVCFYDPEVGNYYGCGHPMKPHRIIMTHALAHYGLDQHMQ 60
 Db 1 mesgnslpdsgdvkrkvefydpevgnyyygqghpmkphritmtallahyglldqhmq 60
 QY 61 VLKFPAREDDICRFHADDDVVSFLKSTTPETQDDQLRLKRFNVGECVPVFGIYSFCQT 120
 Db 61 vlkmaekdrldlckfhaddvafirgltpetqdgirqlkrfnvedcpxfdgysfcqt 120
 QY 121 YAGSGVSGVKLNHGLODDIANNAGLHAKKCEASGFCVNDIVLAILLLQHEHVLV 180
 Db 121 yaagsvsgalklnhgvcdainwaglhakkeasgfcyvndivlailellkhehvlv 180
 QY 181 VDIDHHGDXVEAFAVATDEVMVTSFHKFGDYFPGTGIDIGYSGKYYSLNPLDDGI 240

Db 181 vdidhbgdveaefytdtrvmvtsfhkfgdyfpgtdlridyagkykslnvpiddgi 240

QY 241 DDESYHLKRPIMKGWMEIFRPGAVVLGGADSLSGDLGECNLSTKGAEVCKEKRNFN 300

Db 241 ddesyglfkpimgkwmelfrpgavvlqcgadslsgdrlgcnlstkgaecvkrnf 300

QY 301 VPLLGGGGYTTIRNVARCMCYETGVALGVEEDKKPHEHEYEYEGPDVTLVAPSNNEN 360

Db 301 vpll11999yvtlrnvarcwcetcsvalgllddkmpqheyeiygpdvtlvapsnn 360

QY 361 KNSRQMLEIRNDLHNLKSLQHAPSVPFQERPPDTERPEVEDQEDGKRDPPSDMDV 420

Db 361 knsrqlldelraklldnlslrqlhqvpsvqfgerppdelllerdedqdrderwdpsdrev 420

QY 421 DDDRKPIPSKVRKAEVPEPTKDKGLKGMKGGCEVEVDESGTKVYGVNPGVEAS 480

Db 421 gddsnprvrrrvkscev-daedkd-----tvsqvdsmavdep 456

QY 481 VKMEE 466

Db 457 lkeeqd 462

RESULT 4

AAV28800 standard: Protein: 517 AA.

AAV28800; 13-JAN-2000 (first entry)

XX Maize histone deacetylase-4.

XX Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;

KW acetyl modification; promoter; regulatory element; transgenic plant;

KW disease resistance; toxin screening; pathogenicity;

XX disease response promoter.

OS Zea mays.

PN WO9951731-A2.

PD 14-OCT-1999.

PF 02-APR-1999; 99WO-US07370.

PR 03-APR-1998; 98US-0080563.

PA (PION-) PIONEER HI-BRED INT INC.

PI Baldwin DA, Briggs SP, Crane VC;

DR WPI: 1999-611038/52.

DR N-PSDB; AAX90840.

PT New deacetylase genes, used for producing transgenic plants which have increased disease resistance

PS Claim 1: Page 63-65; 87pp: English.

XX The present sequence is maize histone deacetylase encoded by HD cDNA belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl modifications, may be localized to promoters targeted for repression by other proteins that associate with HD and specifically bind regulatory elements in promoter DNA. The HD nucleotide sequence can be used for CC producing transgenic plants with increased disease resistance. CC Additionally, compositions find use in screening for toxins that affect CC pathogenicity and in determining which disease response promoters are regulated by histone deacetylase.

SQ Sequence 517 AA;


```

XX XX Maize histone deacetylase; HD: HD cDNA; family 1, zmd1; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
XX Zea mays.
XX
XX MO9951731-A2.
XX
XX 14-OCT-1999.
XX
XX 02-APR-1999; 99WO-US07370.
XX
XX 03-APR-1998; 98US-0080563.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Baldwin DA, Briggs SP, Crane VC;
XX
XX WPI: 1999-611038/52.
XX
XX N-PSDB: AAX90839.
XX
XX New deacetylase genes, used for producing transgenic plants which have
XX increased disease resistance
XX
XX Claim 1; Page 58-60; 87pp; English.
XX
XX The present sequence is maize histone deacetylase encoded by HD cDNA
XX belonging to family 1, zmd1. This enzyme responsible for removing acetyl
XX modifications, may be associated to promoters targeted for repression by
XX other proteins that associate with HD and specifically bind regulatory
XX elements in promoter DNA. The HD nucleotide sequence can be used for
XX producing transgenic plants with increased disease resistance.
XX Additionally, compositions find use in screening for toxins that affect
XX pathogenicity and in determining which disease response promoters are
XX regulated by histone deacetylase.
XX
XX Sequence 439 AA:
XX
Query Match 71.7%; Score 1950.5; DB 20; Length 439;
Best Local Similarity 84.2%; Pred. No. 1.2e-194;
Matches 358; Conservative 29; Mismatches 35; Indels 3; Gaps 3;
OY 4 GCNSL-ASGPGGVKRVKVFYDPEYGNYYGGGHPMKRHRIRMTALLAHYGLLOHMOVL 62
DB 8 gnsplrtgagdkrtvcyfydaevgnyygggphmkhrirmtallgryglldqmqvf 67
OY 63 KPRPARDLGRFADADVSLRSTPTODOTROLKRFNVGDCGPFDDGLYSCOTYA 122
DB 68 rphprtdldlctfhaddvstlrsvtpepqqlalrtrnvgedcprfdg1ystcqla 127
OY 123 GGSVGSYKLNHGLDIAINNAGGTHAAKCEASGFCVNDIVLLELQHEHVLVVD 182
DB 128 ggsyvgavklnhgl-hdaiinnagglhhakkeasgfcyvdnlviallellyhqvlyvd 186
OY 183 IDIHGGDEEAPATDTBMTVVSFHKEGDYRGTHIODIGYSGKYYSLNVLDDGIDD 242
DB 187 idihgdeveaftyldtvmvsvfkhfgdyfrgtdirdvghskyslnvpriddgidd 246
OY 243 ESYHLFPIMGKWEIRPGAVVYQCGADSLSDRLGCFNLSTKGAHECYKFRSEFVNP 302
DB 247 esygsllfprimgkwevnpavvayqcgadslsgdr1gcfnlstkghaecvrfmrstnvp 306
OY 303 LLLGGGCTTIRNARFCMCYETGVAGVEVEDKMPHEHYEYFGDYTLHAPSMEKN 362
DB 307 lllggggyttrnarcwcycetgvalshelcdkmpnehyeyfgdytlhapsmekn 366
OY 363 SRQMLEETRNLDLNLKSLQNAAPSVFOERPPDTETPEVDEQDDGCKRMDPDSMDVDD 422
DB 367 trhgldlkskldlslsktrnapsvqfgerppae1pedgedkenpderhdadsdvem-n 425

```

```

OY 423 DRKPI 427
DB 426 dakpi 430
RESULT 7
AAG05187
ID AAG05187 standard; Protein; 471 AA.
AC AAG05187;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1499.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 28-APR-1999; 99US-0130891.
XX
XX 30-APR-1999; 99US-0131449.
XX
XX 30-APR-1999; 99US-0132048.
XX
XX 04-MAY-1999; 99US-0132407.
XX
XX 05-MAY-1999; 99US-0132484.
XX
XX 06-MAY-1999; 99US-0132485.
XX
XX 06-MAY-1999; 99US-0132486.
XX
XX 07-MAY-1999; 99US-0132487.
XX
XX 11-MAY-1999; 99US-0134256.
XX
XX 14-MAY-1999; 99US-0134218.
XX
XX 14-MAY-1999; 99US-0134219.
XX
XX 14-MAY-1999; 99US-0134221.
XX
XX 14-MAY-1999; 99US-0134370.
XX
XX 18-MAY-1999; 99US-0134768.
XX
XX 19-MAY-1999; 99US-0134941.
XX
XX 20-MAY-1999; 99US-0135124.
XX
XX 21-MAY-1999; 99US-0135353.
XX
XX 24-MAY-1999; 99US-0135629.
XX
XX 25-MAY-1999; 99US-0136021.
XX
XX 27-MAY-1999; 99US-0136392.
XX
XX 28-MAY-1999; 99US-0136782.
XX
XX 01-JUN-1999; 99US-0137222.
XX
XX 03-JUN-1999; 99US-0137528.
XX
XX 04-JUN-1999; 99US-0137502.
XX
XX 07-JUN-1999; 99US-0137724.
XX
XX 08-JUN-1999; 99US-0138094.
XX
XX 10-JUN-1999; 99US-0138540.
XX
XX 10-JUN-1999; 99US-0138847.
XX
XX 14-JUN-1999; 99US-0139119.
XX
XX 16-JUN-1999; 99US-0139452.
XX
XX 16-JUN-1999; 99US-0139453.
XX
XX 17-JUN-1999; 99US-0139492.

```


XX 24-AUG-2000; 2000CA-2316036.
 PF
 XX
 XX 27-AUG-1999; 99US-0383971.
 PR
 XX
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 XX
 PI Miki B, Brown D, Tian L, Wu K;
 DR WPI: 2001-258457/27.
 DR N-PSDB: AAF80351.
 XX
 PT Methods for regulating gene expression in transgenic plants, e.g.
 PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
 PT elongation), comprises introducing genes encoding histone deacetylase
 PT
 XX
 PS Claim 10; Fig 1B; 91pp; English.
 XX
 CC The present sequence represents Arabidopsis thaliana histone deacetylase
 CC designated AtRpd3B. The protein is homologous to yeast Rpd3 and HD1.
 CC The polynucleotide sequence is used in the method of the invention.
 CC The specification describes a method for regulating gene expression in
 CC transgenic plants. The method comprises modifying histones by introducing
 CC chimeric nucleotide sequences which have regulatory elements in operative
 CC association with a gene of interest or with a nucleotide sequence
 CC encoding histone deacetylase. The method is useful for regulating the
 CC developmental, physiological or biochemical pathway within a plant,
 CC particularly for repressing ethylene-responsive phenotypes
 CC (e.g. inhibition of hypocotyl elongation). The method is also useful
 CC as a functional test for identifying a phenotype associated with
 CC perturbing a gene. The histone deacetylase genes are useful for
 CC altering the development of an organism.
 XX
 SQ Sequence 471 AA;

Query Match 57.3%; Score 1560.5; DB 22; Length 471;
 Best Local Similarity 60.7%; Pred. No. 7.5e-154;
 Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

QY 2 DTGNSLASGPDGVKRVYFYDEPVGNVYXGQGHPRKPHRIRMTALAHYGLQHMV 61
 DB 4 desglslpsgpdgkrrvsyfyepcltdgyyggqghpmkphrimahsllyhhlrrlel 63
 QY 62 LKRPAREDLCRFHADYVSFLASTIPETQD--QIRQLKRVNGEDCVFGLYSFCQ 119
 DB 64 srpsladasdgtrhspcyvdlasvpsesmgsaaatnrlrrfngedcvfqlfclfor 123
 QY 120 TYAGSGVGSYKLNHGLCDIAINMAGGLHNAKCEASGFCVNDIVLALILKQHRVYL 179
 DB 124 asaggslyaaavklnrqdadalnwg99lnhakkseasgfcvndivlgllelkmfkrvl 183
 QY 180 YVIDIHNGDVEEAFYATDRVMTVSFNKFGDYPFGTGHODIGSGKXYSLNVPRLDDG 239
 DB 184 yldidvhngdveeaftldrvmtvsfnkfgdtpfgtghlrdvgaekgkylahvplndg 243
 QY 240 IDDESYNHLEFRPMKWEIIFRCAVYLQCGAUSLSGDRLCFNLSTKHAECYKFRKF 299
 DB 244 mdesfslsflrpllykymevyqpaavylqcgadslsqdrldcfnlskghadclrlfrrsy 303
 QY 300 NVPLLILAGGCVYTRINVARCWCYETGVALGVEDEKMPENHYEYFPGPDYTLHYAPSNME 359
 DB 304 nvppllmvlggcvytrlnvarcwcylavayvepdklprymeyfeyfpgdytlhvdpspme 363
 QY 360 NKSRLQLEIRNDLNLNLSKLGAPSVPRQERPDTEPEVDEDDQDGDGRMP----- 414
 DB 364 nlneprkmerlrntleqslghapsvqfghtrpvrntv--ldepeddmetrprlrlwsg 421
 QY 415 --DDMDVDDDKRKRIPSRVKAVERPDTKDKGLKIMEKRGKCEVEVDESGSKVTGYN 472
 DB 422 latesdsdddkp-----lhgyscrg-gatldrdsygedemaddn 461

QY 473 P 473
 DB 462 P 462

RESULT 10
 ID AAY28797
 ID AAY28797 standard; protein; 458 AA.
 XX
 AC AAY28797;
 XX
 DT 13-JAN-2000 (first entry)
 XX
 DE Maize histone deacetylase-1.
 XX
 KW Maize histone deacetylase; HD; HD cDNA; family 1, zmHD1; gene repression;
 KW acetyl modification; promoter; regulatory element; transgenic plant;
 KW disease resistance; toxin screening; pathogenicity;
 KM disease response promoter.
 XX
 OS Zea mays.
 XX
 PN W0951731-A2.
 XX
 PD 14-OCT-1999.
 XX
 PF 02-APR-1999; 99MO-US07370.
 XX
 PR 03-APR-1998; 98US-0080563.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Baldwin DA, Briggs SP, Crane VC;
 DR WPI: 1999-611038/52.
 DR N-PSDB: AAX90837.
 XX
 PT New deacetylase genes, used for producing transgenic plants which have
 PT increased disease resistance
 XX
 PS Claim 1; Page 49-51; 87pp; English.
 XX
 CC The present sequence is maize histone deacetylase encoded by HD cDNA
 CC belonging to family 1, zmHD1. This enzyme responsible for removing acetyl
 CC modifications, may be localised to promoters targeted for repression by
 CC other proteins that associate with HD and specifically bind regulatory
 CC elements in promoter DNA. The HD nucleotide sequence can be used for
 CC producing transgenic plants with increased disease resistance.
 CC Additionally, compositions find use in screening for toxins that affect
 CC pathogenicity and in determining which disease response promoters are
 CC regulated by histone deacetylase.
 XX
 SQ Sequence 458 AA;

Query Match 56.9%; Score 1548; DB 20; Length 458;
 Best Local Similarity 63.3%; Pred. No. 1.4e-152;
 Matches 280; Conservative 66; Mismatches 80; Indels 16; Gaps 3;

QY 7 SLASGPDGVKRVYFYDEPVGNVYXGQGHPRKPHRIRMTALAHYGLQHMVLPKP 66
 DB 12 spagsgcdahrrrrsyfyepsltdgyyggqghpmkphrimahslvlyghlrrlelrryp 71
 DB 72 aseadlrrfhsdyvafllaatgpnvylpdralkrrfngedcvfqlfpcqasagysl 131
 QY 67 AREBDCRFHADDYVSFLASTIPETQDQIRQLKRVNGEDCVFGLYSFCQTVAGSV 126
 DB 127 GGSYKLNHGLCDIAINMAGGLHNAKCEASGFCVNDIVLALILKQHRVLYVDIDH 186
 DB 132 gaavklnrqdadltnwag99lnhakkseasgfcvndivlgllelkmfkrvlvldvhl 191
 QY 187 HGDGVEEAFYATDRVMTVSFNKFGDYPFGTGHODIGSGKXYSLNVPRLDDSIDESYH 246
 DB 187 hgdgveeafyatrvmvtsfnkfgdypfgtghlrdvgaekgkylahvplndg 461

PA (IMCO) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Bulwela L, Ali S;
 XX
 DR WPI; 2001-138069/14.
 XX
 XX Suppressing expression of selected gene for treating cancer, involves
 PT introducing peptide comprising nucleic acid binding portion that binds
 PT to or associated with selected gene and chromatin inactivation portion
 PT
 XX Disclosure; Fig 5; 65pp; English.
 XX
 XX The present invention describes a method of suppressing the expression of
 CC a gene of interest by introducing into the cell a protein containing a
 CC DNA binding site and a chromatin inactivation portion, or a nucleotide
 CC encoding such a peptide. Preferably the chromatin inactivation portion is
 CC part of the histone deacetylation (HDAC) complex. The method can be used
 CC in disease treatment, for example in the treatment of cancer by the
 CC suppression of oncogenes, and in the production of disease models.
 XX
 XX Sequence 482 AA:

Query Match 56.7%; Score 1544; DB 22; Length 482;
 Best Local Similarity 59.4%; Pred. No. 4,1e-152;
 Matches 291; Conservative 67; Mismatches 98; Indels 34; Gaps 7;

OY 14 GVRKVCYFDEPVNGYVYGGHPRMHRIMTHALLAHYGLQHMQVLKPPRAREDLIC 73
 DB 6 gtrkvcyydgvgnyyqgphmkphrirmthallahyglrkmelyrphknaeemt 65
 OY 74 RFHADVVSFLRSITPETQODIROLKRFNVGDCVPEDGLYFCQYVAGSGVSKLN 133
 DB 66 kysddylkflrstrpdmseyskqmqtrnvgecdpvtgdlfctqstgsvasavakln 125
 OY 134 HGLCDIAINMAGGLHNAKKEASGFCYVNDIVLAILLELKQHERVLYVDIDIHGDGVEE 193
 DB 126 kgtdlaivnawagjlhakkseasgfcyvndivlaillelkqhgrvlyididihgdgvee 185
 OY 194 AFATDVMVTSFRKRGDVPFGTHIDIGSGSKYSLNVPRLDDGIDDESYNLLFKPRIM 253
 DB 186 atyldtvmvtsfrkgyeyfpigtldiridgaqkkyavnyprldgiddesyeaalfkprv 245
 OY 254 GKWEIEFRPGAVVLOCGASLSGDRIGCFNLSTIKGHAECVKKFMRSPNVLILLGGGGYT 313
 DB 246 skvwmelqpsaavvlgcsdsisgdrigcfnlstikghakcevefksfnlpmimlmgsggyt 305
 OY 314 RNVARCWCYETGVALGVEVEDKMPREHEVEYFGPDYTLHAVSPNMENKSRMLLETRND 373
 DB 306 rnvarcwtetavaldetelpneipyndyfeypdfklhispnmtngntneylekqtr 365
 OY 374 LLAHLSKLOHAPSVFPOERPPDTEPREV--DEODEDGDKRD-----PDSDM 418
 DB 366 lfenlmlpnapvqmqaipedaipseesgededdpdkrtisicssdkrlaaceefsdsee 425
 OY 419 DVDDDRKRPIPS-----RVKREAVEPDTKDKGLKIMERKGCCEVEDESGTKTYGVNP 473
 DB 426 egeegrtvnsnffkakrvkte----dekekd-----evteeektkeeekrea 472
 OY 474 VGVEEASVKM 483
 DB 473 kyvke-evkl 481

RESULT 13
 ID AAB49954 standard; Protein; 482 AA.
 XX AAB49954:
 XX AC AAB49954:
 XX DT 08-MAR-2001 (first entry)
 XX

DE Human histone deacetylase HDAC-1.
 XX
 XX Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
 KM HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;
 KW gene therapy.
 XX
 XX Homo sapiens.
 XX OS
 XX MO200071703-A2.
 PN
 XX
 XX 30-MAY-2000.
 PD
 XX
 XX 03-MAY-2000; 2000MO-1B01252.
 PF
 XX
 XX 03-MAY-1999; 9905-0132287.
 PR
 XX
 XX (METH-) METHYLENE INC.
 PA
 XX
 XX Macleod AR, Li Z, Besterman JM;
 PI
 XX
 XX WPI; 2001-016407/02.
 DR
 XX
 XX N-PSDB; AAC89554.
 DR
 XX
 XX Antisense oligonucleotide that inhibits expression of a histone
 PT deacetylase, useful for treating and/or alleviating the symptoms of
 PT neoplasia, or for inhibiting neoplastic cell growth in an animal -
 PT
 XX Disclosure; Page 49-50; 125pp; English.

PS The present invention provides inhibitors of histone deacetylase enzymes
 CC such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
 CC inhibitors may be antisense strands or they may be compounds identified
 CC by contacting the enzyme with the compound and measuring the resulting
 CC enzyme activity. These inhibitors are useful for treating cancers and for
 CC identifying which histone deacetylase is involved in a neoplasia.
 XX
 XX Sequence 482 AA:

Query Match 56.7%; Score 1544; DB 22; Length 482;
 Best Local Similarity 59.4%; Pred. No. 4,1e-152;
 Matches 291; Conservative 67; Mismatches 98; Indels 34; Gaps 7;

OY 14 GVRKVCYFDEPVNGYVYGGHPRMHRIMTHALLAHYGLQHMQVLKPPRAREDLIC 73
 DB 6 gtrkvcyydgvgnyyqgphmkphrirmthallahyglrkmelyrphknaeemt 65
 OY 74 RFHADVVSFLRSITPETQODIROLKRFNVGDCVPEDGLYFCQYVAGSGVSKLN 133
 DB 66 kysddylkflrstrpdmseyskqmqtrnvgecdpvtgdlfctqstgsvasavakln 125
 OY 134 HGLCDIAINMAGGLHNAKKEASGFCYVNDIVLAILLELKQHERVLYVDIDIHGDGVEE 193
 DB 126 kgtdlaivnawagjlhakkseasgfcyvndivlaillelkqhgrvlyididihgdgvee 185
 OY 194 AFATDVMVTSFRKRGDVPFGTHIDIGSGSKYSLNVPRLDDGIDDESYNLLFKPRIM 253
 DB 186 atyldtvmvtsfrkgyeyfpigtldiridgaqkkyavnyprldgiddesyeaalfkprv 245
 OY 254 GKWEIEFRPGAVVLOCGASLSGDRIGCFNLSTIKGHAECVKKFMRSPNVLILLGGGGYT 313
 DB 246 skvwmelqpsaavvlgcsdsisgdrigcfnlstikghakcevefksfnlpmimlmgsggyt 305
 OY 314 RNVARCWCYETGVALGVEVEDKMPREHEVEYFGPDYTLHAVSPNMENKSRMLLETRND 373
 DB 306 rnvarcwtetavaldetelpneipyndyfeypdfklhispnmtngntneylekqtr 365
 OY 374 LLAHLSKLOHAPSVFPOERPPDTEPREV--DEODEDGDKRD-----PDSDM 418
 DB 366 lfenlmlpnapvqmqaipedaipseesgededdpdkrtisicssdkrlaaceefsdsee 425
 OY 419 DVDDDRKRPIPS-----RVKREAVEPDTKDKGLKIMERKGCCEVEDESGTKTYGVNP 473
 DB 426 egeegrtvnsnffkakrvkte----dekekd-----evteeektkeeekrea 472
 OY 474 VGVEEASVKM 483
 DB 473 kyvke-evkl 481

Db	426	egegirkusnmtfkakrvkte----	dekexd-----	peekK-----	evleeeektkeepka	472
Oy	474	WGEESAYKM	483		:	
Db	473	kayke-evkl	481			
RESULT 14						
ID	AAB56985					
XX	AAB56985 standard; Protein: 488 AA.					
AC	AAB56985;					
XX	13-MAR-2001 (first entry)					
DT						
DE	Human prostate cancer antigen protein sequence SEQ ID NO:1563.					
XX						
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;					
KW	neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;					
KW	vulnerable; gastrointestinal; nephrotropic; antiinfective; gynaecological;					
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;					
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;					
KW	wound; infectious disease.					
XX						
XX	Homo sapiens.					
OS						
PN	MO2005174-AI.					
XX						
PD	21-SEP-2000.					
XX						
PF	08-MAR-2000; 2000MO-US05988.					
PR	12-MAR-1999; 99US-0124270.					
XX						
PA	(HUMA-) HUMAN GENOME SCI INC.					
XX	(ROSE/) ROSEN C A.					
PI	Rosen CA, Ruben SM;					
DR	WPI: 2000-587513/55.					
XX	N-PSDB; AAF16188.					
PT	Prostate cancer associated gene sequences, referred to as prostate					
PT	cancer antigens, useful for treatment, prevention, and diagnosis of					
PT	disorders such as prostate cancer -					
PX	Clam 11; Page 2005-2006; 233pp; English.					
XX						
CC	AAF15566 to AAF16505 encode the human prostate cancer associated					
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.					
CC	The prostate cancer antigens can have neuroprotective, cytosolic,					
CC	cardioscopic, immunomodulatory, muscular, vulnery, gastrointestinal,					
CC	nephrotropic, antiinfective, gynaecological and antibacterial activities,					
CC	and can be used in gene therapy. The prostate cancer antigen					
CC	polynucleotides may be used for detection of prostate cancer, chromosome					
CC	identification, as chromosome markers, and for numerous other diagnostic					
CC	or research purposes. The prostate cancer antigens may be used to treat					
CC	disorders such as neural, immune, muscular, reproductive,					
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative					
CC	disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to					
CC	AAB57303 represent sequences used in the exemplification of the present					
CC	Invention.					
XX						
SQ	Sequence 488 AA:					
Query Match 56.7%; Score 1544; DB 21; Length 488;						
Best Local Similarity 59.4%; Pred. No. 4.2e-152;						
Matches 291; Conservative 67; Mismatches 98; Indels 34; Gaps 7;						
OY	14	GKKRVCYFYDDEVGNCNYYGGCGHPKPHRIKTHALHLAHYGLLOHMNOVLKPPAPAEERDLG	73		:	:
DB	12	gttttccyyggddvnyyyggghpmkphrhmrlmllllygltymelytpkhanaeent	71	:	:	:

```

QY 74 RHMADYVSFRLSTPEFOQDOIROLKRFVNWGEDCPVEDGLYSFCQTVAGSGVSGSYKLN 133
Db 72 kyhdbdykfkfrstirpdmseysskqmgrfvngedcprfdglfetcqslsgsvsasavkln 131
QY 134 HGLCDLAINMAGSLHNAKKEASFCYVNDVLAILELLKQHEBRLVYVNDIHNGXQVEE 193
Db 132 kqgdaiannwagsgjhakkseasfeyvndviallailkyhgvrlyididhngdvee 191
QY 194 AFVATDRMTNMFPIKPDQYPRGTGTHIODIGVSGSKYVSLNWPRDDGIDDESSYHLNFKRPM 253
Db 192 afyctdrmtvmsfthkyeyfprgtdrltdagkkyavnyprltdgddesyeaifkrpm 251
QY 254 GKWMEIFRPGAVNLQCGADSLSSGRLCCFNLISKGHAECVAFMKSFWNPPLLGGGGYTL 313
Db 252 skvmemfgrsavvlyqgsdsjsgdrlgcnlltkghakvefvsksfplmlm199gytll 311
QY 314 RNVARCMCYETFGVALGYEVEDEKMEHEHEYTEFGVDYTLHAAPSMEKNSKOMLEELRND 373
Db 312 rnvarcmyetfvalgyeivdeipneirlyndyfeyfgrdflhpsmtnlqnleneylklqr 371
QY 374 LLHNLSTKIOHAPSVPRFOERPRDPTREY-DEDDODGCKRND-----PDSDM 418
Db 372 lfennlphrpyvqmqaipeasgdeddddpktrrisscdkrrisceeefsdssee 411
QY 419 DVDDDRKRPIS-----RYKREAVEBPTRKDKDGLKGIMEKRGCEVEYDESGSTRVTCVNP 473
Db 432 egeggrfkrsnffkkrakrvkte-----dekek-----evteeektkkeeprea 478
QY 474 VGVEEASVXKM 483
Db 479 kvvke-evkl 487

```

CC	AA088919	is a transcriptional control protein, RPDL, derived from a human foetal lung cDNA library The gene encoding RPDL is localised
XX	AA088919	
XX	AA088919	standard; Protein; 482 AA.
XX	AA088919;	
XX	09-SEP-1996	(first entry)
XX	RPDL, a human foetal lung-derived transcriptional control protein.	
XX	Transcription; expression; control; diagnosis; study; cancer;	
XX	mammary; gastric carcinoma; chromosome 1p34.1.	
XX	Homo sapiens.	
XX	EP708112-A1.	
XX	24-APR-1996.	
XX	21-SEP-1995;	95SEP-0114884.
XX	22-SEP-1994;	94UP-0227876.
XX	(CANC-) CANCER INST.	
XX	(EISA) EISAI CO LTD.	
XX	Furukawa Y, Nakamura Y;	
XX	WPI: 1996-202003/21.	
XX	N-PSDB; AAT12940.	
XX	New human transcriptional control protein RPDL - used to develop	
XX	prods. for study and diagnosis involving the protein, partic. for	
XX	gene analysis	
XX	Claim 1; Page 10-12; 15pp; English.	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:57:27 ; Search time 44.51 Seconds

(Without alignments)
1547.838 Million cell updates/sec

Title: US-09-645-337-4

Perfect score: 2544

Sequence: 1 MEADSGISLPSGPDGKRRR.....TGEDEMDNDPPDVNPSS 471

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: SP_Archaea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Mammal:*
8: SP_Mhc:*
9: SP_Organelle:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2542	99.9	471	10 Q9FVE5	Q9FVE5 arabidopsis
2	2535	99.6	471	10 Q9FVL2	Q9FVL2 arabidopsis
3	1738	68.3	458	10 Q9ZTP8	Q9ZTP8 zea mays (m
4	1560.5	61.3	501	10 Q9SLZ3	Q9SLZ3 arabidopsis
5	1522	59.8	501	10 Q9LKG1	Q9LKG1 mesembryant
6	1496.5	58.8	493	10 Q9AXF0	Q9AXF0 oryza sativ
7	1476	58.0	521	5 Q77213	Q77213 drosophila
8	1476	58.0	521	5 Q9VZA1	Q9VZA1 drosophila
9	1373	54.0	687	3 Q9PAF5	Q9PAF5 drosophila
10	1369	53.8	409	10 Q9FHO9	Q9FHO9 arabidopsis
11	1305.5	51.3	465	5 Q62339	Q62339 caenorhabdl
12	1292.5	50.8	428	11 Q99PA0	Q99PA0 rattus norv
13	1288	50.6	444	5 Q9GUA8	Q9GUA8 cryptospori
14	1284	50.5	648	3 Q9CIC6	Q9CIC6 cochicobol
15	1281	50.4	448	5 Q9XVC7	Q9XVC7 plasmodium
16	1277.5	50.2	428	11 Q9UM08	Q9UM08 mus muscula
17	1275.5	50.1	428	11 Q9JLX5	Q9JLX5 mus muscula
18	1275	50.1	405	3 Q59702	Q59702 schizosacch
19	1233.5	48.5	438	5 Q9VNC2	Q9VNC2 drosophila

20	1229	48.3	566	3 Q9HDT2	Q9HDT2 ustilago ma
21	1207	47.4	419	10 Q9M1N6	Q9M1N6 arabidopsis
22	1151.5	45.3	437	5 Q9GUS9	Q9GUS9 cryptospori
23	1072	42.1	429	5 Q9GTF4	Q9GTF4 tetrahymena
24	994.5	39.1	481	3 Q9P4F4	Q9P4F4 emericella
25	820.5	32.3	377	11 Q9DOK6	Q9DOK6 mus musculu
26	813	32.0	377	4 Q9NE76	Q9NE76 homo sapien
27	808	31.8	377	4 Q9NH44	Q9NH44 homo sapien
28	748.5	29.4	428	5 Q9GRP1	Q9GRP1 leishmania
29	596.5	23.4	256	11 Q99PA2	Q99PA2 rattus norv
30	586	23.0	223	4 Q9H368	Q9H368 homo sapien
31	457.5	18.0	375	2 Q67135	Q67135 aquifex aeo
32	424.5	16.7	367	2 Q9WX04	Q9WX04 streptomyce
33	410.5	16.1	389	2 Q9K7X1	Q9K7X1 bacillus ha
34	371.5	14.6	389	2 Q9GTC9	Q9GTC9 staphylococ
35	319.5	12.6	380	2 Q9HXM1	Q9HXM1 pseudomonas
36	305	12.0	310	2 Q67877	Q67877 aquifex aeo
37	299.5	11.8	158	10 Q9LXN8	Q9LXN8 arabidopsis
38	290.5	11.4	883	5 Q9XYX1	Q9XYX1 drosophila
39	277	10.9	1108	10 Q9FNO7	Q9FNO7 arabidopsis
40	273.5	10.8	1114	11 Q9JL73	Q9JL73 mus musculu
41	269.5	10.6	425	10 Q9S3E6	Q9S3E6 arabidopsis
42	269.5	10.6	796	5 Q17323	Q17323 caenorhabdl
43	268	10.5	338	1 Q9HSP7	Q9HSP7 halobacteri
44	266.5	10.5	577	10 Q9LS38	Q9LS38 arabidopsis
45	265	10.4	878	4 Q9URU7	Q9URU7 homo sapien

ALIGNMENTS

RESULT 1

Q9FVE5 PRELIMINARY; PRT; 471 AA.

AC Q9FVE5; 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE PUTATIVE HISTONE DEACETYLASE.

GN RPD3B.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI-TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Wu K., Malik K., Tian L., Brown D., Miki B.;

RT *Functional analysis of RPD3 histone deacetylase homologs in

Arabidopsis thaliana.*

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF195548; AAC28475.1; -

DR InterPro: IPR000286; His_deacetylase.

DR Pfam: PF00850; Hist_deacetyl1.1.

DR PRINTS: PR01270; HDAUPER.

SO SEQUENCE 471 AA; 52720 MW; 371BF7040E508649 CRC64;

Query Match 99.9%; Score 2542; DB 10; Length 471;

Best Local Similarity 100.0%; Pred. No. 5.3e-198;

Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEADSGISLPSGPDGKRRVSYFEPTIGDYGGCHPMKPHRIRMAHSLIIHYHLARR	60
QY	1	MEADSGISLPSGPDGKRRVSYFEPTIGDYGGCHPMKPHRIRMAHSLIIHYHLARR	60
QY	1	MEADSGISLPSGPDGKRRVSYFEPTIGDYGGCHPMKPHRIRMAHSLIIHYHLARR	60
QY	61	LEIRSPSLAASDGRHSPHYVDFLASVSPESMGDSAAARNLRFRVNGEDCPYFOLF	120
QY	61	LEIRSPSLAASDGRHSPHYVDFLASVSPESMGDSAAARNLRFRVNGEDCPYFOLF	120
QY	61	LEIRSPSLAASDGRHSPHYVDFLASVSPESMGDSAAARNLRFRVNGEDCPYFOLF	120
QY	121	FCRASAGSICAAVKLRPADADIAINMGCGLHRAKKGEASGFCVNVIVIGITELLMFK	180
QY	121	FCRASAGSICAAVKLRPADADIAINMGCGLHRAKKGEASGFCVNVIVIGITELLMFK	180
QY	121	FCRASAGSICAAVKLRPADADIAINMGCGLHRAKKGEASGFCVNVIVIGITELLMFK	180

Db 121 FCRAAGSISGAAYKLNRODADIAINMGGLHHAKKSPASFCYVNDIVLIGLELHKMK 180
 QY 181 RLVYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYALNPL 240
 Db 181 RLVYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYALNPL 240
 QY 241 NDGMDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHADCLREFL 300
 Db 241 NDGMDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHADCLREFL 300
 QY 301 RSYNPLAVLGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTLHNDPS 360
 Db 301 RSYNPLAVLGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTLHNDPS 360
 QY 361 PMENLNTPKDMERIRNTLLEQLSLIHAPSVQFOHTPPVNRVLDEPEDDMETRPKPRXMS 420
 Db 361 PMENLNTPKDMERIRNTLLEQLSLIHAPSVQFOHTPPVNRVLDEPEDDMETRPKPRXMS 420
 QY 421 GTATYESDSDDDOKPLHGYSRCGATTDSDSTGEDEMDDDNPEPDVNPSS 471
 Db 421 GTATYESDSDDDOKPLHGYSRCGATTDSDSTGEDEMDDDNPEPDVNPSS 471

RESULT 2

Q9FML2 PRELIMINARY: PRT: 471 AA.
 ID Q9FML2:
 AC 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HISTONE DEACETYLASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA.
 RX MEDLINE:98162728; PubMed:9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones.";
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB008265; BAB10553.1;
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl1;
 DR PRINTS: PR01270; HDASUPER.
 SQ SEQUENCE 471 AA: 52651 MW: CA16C2640DIB1732 CMC64;

Query Match 99.6%; Score 2535; DB 10; Length 471;
 Best Local Similarity 99.6%; Pred. No. 1.9e-197;
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEADSGISLSPGRCRRVSYFEPTIGDYVYGGCHPMKPHIRMAHSLIHHLRR 60
 Db 1 MEADSGISLSPGRCRRVSYFEPTIGDYVYGGCHPMKPHIRMAHSLIHHLRR 60
 QY 61 LEISRPISLADASDGRFSPPEYVDFLASVSPESMGDPSAARLRRFNVGDCPVFDELTD 120
 Db 61 LEISRPISLADASDGRFSPPEYVDFLASVSPESMGDPSAARLRRFNVGDCPVFDELTD 120
 QY 121 FCRAAGSISGAAYKLNRODADIAINMGGLHHAKKSPASFCYVNDIVLIGLELHKMK 180
 Db 121 FCRAAGSISGAAYKLNRODADIAINMGGLHHAKKSPASFCYVNDIVLIGLELHKMK 180
 QY 181 RLVYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYALNPL 240
 Db 181 RLVYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYALNPL 240

QY 241 NDGMDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHADCLREFL 300
 Db 241 NDGMDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHADCLREFL 300
 QY 301 RSYNPLAVLGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTLHNDPS 360
 Db 301 RSYNPLAVLGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTLHNDPS 360
 QY 361 PMENLNTPKDMERIRNTLLEQLSLIHAPSVQFOHTPPVNRVLDEPEDDMETRPKPRXMS 420
 Db 361 PMENLNTPKDMERIRNTLLEQLSLIHAPSVQFOHTPPVNRVLDEPEDDMETRPKPRXMS 420
 QY 421 GTATYESDSDDDOKPLHGYSRCGATTDSDSTGEDEMDDDNPEPDVNPSS 471
 Db 421 GTATYESDSDDDOKPLHGYSRCGATTDSDSTGEDEMDDDNPEPDVNPSS 471

RESULT 3

Q9ZTP8 PRELIMINARY: PRT: 458 AA.
 ID Q9ZTP8:
 AC 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HISTONE DEACETYLASE.
 GN HD1B.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. CUZCO;
 RA Pipal A., Wegener S.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF045473; AAD10139.1;
 DR Mendel: 39025; Zeama; 3043; 39025.
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl1;
 DR PRINTS: PR01270; HDASUPER.
 SQ SEQUENCE 458 AA: 50940 MW: A62775068225BE79 CRC64;

Query Match 68.3%; Score 1738; DB 10; Length 458;
 Best Local Similarity 69.3%; Pred. No. 8.5e-133;
 Matches 328; Conservative 51; Mismatches 72; Indels 22; Gaps 7;

QY 1 MEADSGISLSPGRCRRVSYFEPTIGDYVYGGCHPMKPHIRMAHSLIHHLRR 57
 Db 1 MEADSGISLSPGRCRRVSYFEPTIGDYVYGGCHPMKPHIRMAHSLIHHLRR 57
 QY 58 HRRLISRPISLADASDGRFSPPEYVDFLASVSPESMGDPSAARLRRFNVGDCPVF 115
 Db 58 HRRLISRPISLADASDGRFSPPEYVDFLASVSPESMGDPSAARLRRFNVGDCPVF 115
 QY 116 DGLFPCRAAGSISGAAYKLNRODADIAINMGGLHHAKKSPASFCYVNDIVLIGLEL 175
 Db 116 DGLFPCRAAGSISGAAYKLNRODADIAINMGGLHHAKKSPASFCYVNDIVLIGLEL 175
 QY 176 LKMFRRVLYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYA 235
 Db 176 LKMFRRVLYIDIDVHHGDCGEAEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEKGYA 235
 QY 236 LNVPLNGMDDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHAD 295
 Db 236 LNVPLNGMDDESFRLSPRLIOKWEVYQPEAVVLOCGADSLSGDRLCGPNLSYKSHAD 295
 QY 296 CLRLRSYNNVPLVGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTL 355
 Db 296 CLRLRSYNNVPLVGGEGYTTIRNARCWCYETAAGVGPDKLPLYNNEFEYFGDPDTL 355
 QY 356 HVDPSPMENLNTPKDMERIRNTLLEQLSLIHAPSVQFOHTPPVNRVLDEPEDDMETRP- 414

Db 360 KNSRPMILDDIRGLLESLSLQHPSPVQPERPEAEIPEDEDDHDEPRMDP----- 414
 QY 422 TATESDDDDDKPLHGSCGATTRDSTGDEMDDDNEPEVNP 469
 Db 415 -----DSIDMEVDHDSI-----SIR--SRVKEIMEPRDKRDVDKAVDHP 453

RESULT 6

Q9AXF0 PRELIMINARY; PRT: 493 AA.

AC Q9AXF0:
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE HISTONE DEACETYLASE HDL.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzae; Oryza.
 NCBI_TaxID=4530;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CY. YUANFENGZAO;
 RA Song F., Goodman R.M.;
 RT "Molecular characterization of a rice histone deacetylase gene
 OSHD1.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF332875; AAC01712.1; -;
 SQ SEQUENCE 493 AA; 55054 MW; 0BE300B42243C13 CRC64;

Query Match 58.8%; Score 1496.5; DB 10; Length 493;
 Best local similarity 60.5%; Pred. No. 3,7e-113;
 Matches 294; Conservative 64; Mismatches 97; Indels 31; Gaps 11;

QY 3 ADEGSLRP-SGPRGRKRVSYFEPTIGDYTGCGHPMKPHIRMAHSLIHHLRL 61
 Db 5 AGGGNSLPTAGAGCARRKYCHFDAGVNYCCGHPMKPHIRMTALLAHGLDGM 64
 QY 62 EISRPISLADSDIGRISPEYVDVLASVSPESMGDPASARNLRPNVGEDECPVDFGLF 121
 Db 65 QYLRPHARBDLCRFNADVDVAFRLRSYTPETQD--QIALKRFNNGEDCPVDFGLXSF 122
 QY 122 CRASAGSGAIVKLNQDADIALNMGGLHNAKSEASGFCYVNDIVLGLLELKMFKR 181
 Db 123 CQTYAGSGVGAVKLN-HGHDIALINMAGGLHNAKSEASGFCYVNDIVLGLLELKYQR 181
 QY 182 VLYIDIVNHGDSVEAFRTDRTVTSFHKFGDPFPGTGHIRVGAEKGYALANPLN 241
 Db 182 VLYVIDIDHHDGVEAFRTDRTVTSFHKFGDYFPGTGDIRDIGSEGYGCLNPLD 241
 QY 242 DGMDESEFSLRPLQKVMVEVYQPEAVVLQCGADSLSGDLGCFNLVSKHADCLRFLR 301
 Db 242 DGIDDDVSGISFKPLISKVMEMVRPAGVVIQCGADSLSGDLGCFNLVSKHADCLRFLR 301
 QY 302 SYNPLMLVSGEGYTIRNVAKCMCYETAAVAVGEPDNKLPYNEYFEYFGPDYTLVDPSP 361
 Db 302 SFNPLLLGSGGYAIRNVAKCMCYETAAVAVGELHGLTDKMPNEYFEYFGPREYSLFVAASN 361
 QY 362 MENLTPROMRIRNTLLLEOLSGLIHAPSVQFHTPPYNNKRLVDEPEDMETRPKRWMSG 421
 Db 362 MENNTNMQLEIKNLTDLNLSKLDHAPSVQFEEIRIPETK-LPEPDEDDQED-PDER---- 415
 QY 422 TATESDSD--DDDKPLHGSCG--GATTRDSTGDEMDDDN-----PEPD 465
 Db 416 ---HDPDSQMKVLDHKKPR-GHSARSLIRNLEVKREIT-ESKAKDHGKRLTTEHKGFEP 470
 QY 466 VNPSS 471
 Db 471 ADUPGS 476

RESULT 7

077213 PRELIMINARY; PRT: 521 AA.

AC 077213:
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DE PUTATIVE HISTONE DEACETYLASE.
 GN RPD3 OR CG7471.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Mottus R.C., Sobel R.E., Grigliatti T.A.;
 RT "DMHDA1";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF086715; AAC61494.1; -;
 DR Flybase: FBgn0015805; Rpd3.
 DR InterPro: IPR00286; His.deacetylase.
 DR Pfam: PF00850; Hist.deacetyl. 1.
 DR PRINTS: PR01270; HDASUPER.
 SQ SEQUENCE 521 AA; 58344 MW; B0F6503D42A1BA32 CRC64;

Query Match 58.0%; Score 1476; DB 5; Length 521;
 Best local similarity 58.6%; Pred. No. 1.8e-111;
 Matches 273; Conservative 76; Mismatches 89; Indels 26; Gaps 5;

QY 18 KRVSYFEPTIGDYTGCGHPMKPHIRMAHSLIHHLRLISRPSLADSDIGR 77
 Db 6 KRVSYFEPTIGDYTGCGHPMKPHIRMTALLAHGLDGM 64
 QY 78 HSEPYVDVLASVSPESMGDPASARNLRPNVGEDECPVDFGLFPCFASAGSGAIVKLN 137
 Db 66 HSEPYVFLRSIRPDNMEYN--KOMQRFNNGEDCPVDFGLFPCFASAGSGAIVKLN 123
 QY 138 RODADIALNMGGLHNAKSEASGFCYVNDIVLGLLELKMFKRVLYIDIVNHGDSVE 197
 Db 124 KQASEICINMGGLHNAKSEASGFCYVNDIVLGLLELKHVRVLYIDIVNHGDSVE 183
 QY 198 AFYTTDRVMTVSFHKFGDPFPGTGHIRVGAEKGYALANPLNDGMDSEFSLRPL 257
 Db 184 AFYTTDRVMTVSFHKFGDPFPGTGHIRVGAEKGYALANPLNDGMDSEFSLRPL 243
 QY 258 OKYMEVYQPEAVVLQCGADSLSGDLGCFNLVSKHADCLRFLRSYNNPLVSGEGYT 317
 Db 244 SKYMETFQPAVAVVLQCGADSLSGDLGCFNLVSKHADCLRFLRSYNNPLVSGEGYT 303
 QY 318 RNVAKCMCYETAAVAVGEPDNKLPYNEYFEYFGPDYTLVAVPSPMNLTPKDMERIRNT 377
 Db 304 RNVSRCMTVETSVAAVAVGEPDNKLPYNEYFEYFGPDYTLVAVPSPMNLTPKDMERIRNT 363
 QY 378 LLEOLSGLIHAPSVQFHTPPYNNKRLVDEPEDMETRPKRWMSG 425
 Db 364 LFNELRLPLHAPSVQFHTPPYNNKRLVDEPEDMETRPKRWMSG 417
 QY 426 ESDSDDDDKPLHGSCGATTRDSTGDEMDDDNPEPDVNPSS 471
 Db 418 YSDSEDEG-----GGRDRNRSYKQGRKRPDLKDNNSKASS 455

RESULT 8

Q9VZAI PRELIMINARY; PRT: 521 AA.

AC Q9VZAI:
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE RPD3 PROTEIN.
 GN RPD3 OR CG7471.

Db	301	CVKVKVSKENLPTLLIVGGGGYTHRNVARIMAFETCLIVGDNLGSELPYNDYKEKFAPIYEL	360
QY	356	HYDVSPEHNLNTPKPMDEMRITLTLEQLSHAPSVQFHTPP---VNRVLDEPE---ID	409
Db	361	DVRSNMNMNATREYLRITQVYENLRKTFAPSVQMTDVPREPLVDGMDEAEALDD	420
QY	410	M-ETRPKRKXMSGTATYE-----SDSDDD-----KELH-----	437
Db	421	LDEDEKCKKRTKTRKFDQVYEKPGELSDSEDEDEENAAVYTRKPAHLKRHNQANYRLDLA	480
QY	438	--GYSRCGATTDORDSTGEEDD-----DNEPED-----VNPPS 470	
Db	481	DSGVSGMATTQDASSVADEEMDGTGYKLTFAAGCPFPDSSAQTSSAAEPSS 533	

RESULT	10
Q9FH09	
ID	Q9FH09
PRELIMINARY:	
PRT:	409 AA.
AC	Q9FH09;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	HISTONE DEACETYLASE.
OC	Arbidopsins thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;
OC	eurosidis II; Brassicales; Brassicaceae; Arabidopsins.
OX	NCBI_TaxID=3702;
LN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-COLUMBIA:
XX	MEDLINE=20181125; PubMed=10718197;
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA	Tabata S.;
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequences
RT	features of the regions of 3,076,755 bp covered by sixty P1 and PAC
RT	clones.";
RL	DNA Res. 7:31-63(2000).
DR	EMBL; AB023031; BAB09994.1; -;
DR	InterPro; IPR000286; H1s_deacetylase.
DR	Pfam; PF00850; Hist_deacetyl_1.
DR	PRINTS; PR01270; HDASUPER.
SO	SEQUENCE 409 AA; 46017 MW; 56A4FFAF9A0734AE CRC64;

Query Match	53.88;	Score 1369;	DB 10;	Length 409;
Best Local Similarity	63.98;	Pred. No. 6.2e-103;		
Matches 260; Conservative	50;	Mismatches 89;	Indels 8;	Gaps 3

QY	15	DGKRKRVSYFEPTIGDYVYCGOGRPMKPHRIIRMAHSIILHHLRLRLEISPSLADSDI	74
Db	6	DGGRKRVSYFETPMIGDYVYVNOPTKFORIRVHNLHLSLIRHMEINHPDLADSDF	65
QY	75	GRFHSPEVVDLFLASYPESMGD--PSAARNLRPNVGHDC--PVPDFDLFCGRASAGSI	130
Db	66	EKFHSLEIINFLKSTPETHVDPRHPSVSENLRKRNVDYDMGPRFHNLFDTCCRAYAGSI	129
QY	131	GAAYKLNRODADIALNMGGILHNAKSEASGECYVNDIVGLILELMEKRYLIDIVH	190
Db	126	SAAALINFGADIALINMAGGHHNKKOKASGCVYNDIVLAILLELKSFRKLYITEIGP	185
QY	191	HGDGYEEAFYTTDRVMYVSFKHFDPPFGTGHIDVGAKEGKYIALVPLNDGMDDESFR	250
Db	186	HGDEVEEAFKCTDRMYTSSFKHVD---TGDISDYCEGKKQYVSLNAPLKDGDIDFSLR	241
QY	251	SLFRLLQKWMVEYQPRVAVVYQCCADSLISGRLOCFNLVYVGHADCLRLRSTVWPLAVL	310
Db	242	GLFIPVIRHRAEITEPEYIVLQCCADSLIAGDPFETFNLSIKGHDCDLQYVSNFVWPLIL	301
QY	311	GGEGEITINVAKCMCYEAAVAVGEPRDKLBYNFEYFEGGDTYFLHYDSEPMENLNPDK	370
Db	302	GGGGYTLPNVAKCMCYETALVAVGEGLDDPGNDIMKIFRPDYKLHLIPLTNKQMLNRLO	361

QY 371 MERIRNTLLFOLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKR 417
: : | | | | : | | | | : : | | | | |
Db 362 IITMRETLAQLSLVNHAPSVFPQDTPSSSQATEAAEVDMEKRNIDPR 408

```

RESULT 11
062339
ID 062339 PRELIMINARY; PRT: 465 AA.
AC 062339; 062343;
DT 01-APR-1998 (TrEMBLrel. 07, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE R06C1.1 PROTEIN.
GN R06C1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

```

RP SEQUENCE FROM N.A.
RA Kershaw J.²
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
RN (2)

RP SEQUENCE FROM N.A.
RX MEDLINE-94150718: Pubmed-7906398:
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,
RA Lightning J., Lloyd C., Murray A., Mortmore B., O'Callaghan M.,
RA Parsons J., Percy C., Rikken L., Koopa A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
BT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RL Nature 366:32-38(1994).
 DR EMBL, Z81108, CAB03240.1, -.
 DR EMBL, Z81105, CAB03240.1, JOINED.
 DR EMBL, Z81106, CAB03224.1, -.
 DR EMBL, Z81108, CAB03224.1, JOINED.
 DR InterPro: IPR000286; His_deacylase.
 DR Pfam: PF00850; His_deacylase_1.
 DR PRINTS: PR01270; HDASUPER.
 SC SEQUENCE 465 AA: 5270 MW: 325E7ID5197012C5 CRC64:

Query Match	51.38	Score	1305.5	DB	5	Length	465
Best Local Similarity	61.98	Pred. No.	1.1e-97				
Matches	234	Conservative	67	Mismatches	74	Indels	3
						Gaps	2

QY	18	KRVSVEYEPYIGVYVYGGGHPMKPHIIRMAHSIIHYHILHRLREISNPSIADSDIGRE	77
Db	7	KSRVSYYIDGDFGNRYVYGGGHPMKPHYRBMTHSLIVYGLYRKLVNRRAPASPSLTRY	66
QY	78	HSPEYVDFLASVSPESMGDPSAARNLRFPNVGECRPVFDGLFDFCSKASGSGIQAANKLN	137
Db	67	HSDDYINLRNWKSDNMS--FTDOMARFSYGECCPVPFDGMYEFCQISGSGSLAAARNLN	124
QY	138	QODADIAIMNGGGLHNAKSEASGFCYVNDVLTLELLMKFKVLYTODIDVNHGDVEE	197
Db	125	QOESDIAIMNGGGLHNAKSSASGFCYSNDVLTALLELLHNHKKVLYTODIDVNHGDVEE	184
QY	198	AFYTTDRMYTVSPFKFGDFPFGTGHIDVGAEKQKYTALNPLMDGMDSEFNSLPFLI	257
Db	185	AFYTTDRMYTVSPFKHGEYFPGTDLKDVGASQKYTALNPLMDGVDTYERLPIFTIM	244
QY	258	QKWEVYVPEALVYVQCGGDSISGGRGCFNLSYGHADCLRELSYVAVPLVYMGCEYTI	317
Db	245	GEVWARQPEALVYVQCGGDSISAGRLGCVFNLTYYGHGKCYEYMKSPFNVLIIYOGGCTI	304
QY	318	RNVARCWCEYETAVAAGVEPDKLIPNEYFEYFGDVTYLHVDP--SPMENLNTPKIMEHIN	376

QY 124 ASAGSGTCAAVKLNRODADIAIMGGGLHAAKSEASGFCYVNDIVLGLLELLMKFKRL 183
 Db 120 TYAGSGVSGSXLNGLCDIAIMAGGGLHAAKSEASGFCYVNDIVLGLLELLMKFKRL 179
 QY 184 YVIDVHGGDVEEAFYTTDRVMTVSFHKFGDPGTGTHIDVGAKEKKYALNVPRLNDG 243
 Db 180 YVIDIDIHGGDVEEAFYTTDRVMTVSFHKFGDPGTGTHIDVGAKEKKYALNVPRLNDG 239
 QY 244 MDESFRLSFLPRLIOKMEVYQPEAVNLOCGADSLSGRLGCFNLVSGHACDLRELSY 303
 Db 240 IDDESYGLHLLFPIKMKVMEIFRPGAVNLOCGADSLSGRLGCFNLVSGHACDLRELSY 299
 QY 304 NVPLNVLGEGYTTIRNVARCWCYETAFAVGYEPDNKLPYNEFEYFGPDYTLHVPSPME 363
 Db 300 NVPLLLGLGGGYTTIRNVARCWCYETAFAVGYEPDNKLPYNEFEYFGPDYTLHVPSPME 359
 QY 364 NLNTPKDMERIRNTLLEQSLGILHAPSVQFOHTPPVNV--LDEPEDMETRPKPRXMSG 421
 Db 360 NKNRQMLLEIRNDLHNLKLOHAPSVFQERPVTEPREVEDQEDGDRMPDPS-- 416
 QY 422 TATYESDDDDKRP-----LHGYSGRG-CATTDHSTGDEMDDN 461
 Db 417 ----DMVDDDRKRPIPSNVKREAVEPDTKDKDGLKIMERGKCEVEVEDSGSKVTGVN 472
 QY 462 P 462
 Db 473 P 473
 RESULT 2
 HDAC_MAIZE STANDARD; PRT; 513 AA.
 AC P56521;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE HISTONE DEACETYLASE (RPO3 HOMOLOG).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 NC NCHI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. W22;
 RA Rossi V., Hartings H., Motto M.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF035815; AAC50038.1; -
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl1; 1.
 DR PRINTS: PRO1270; HDASUPER.
 DR PRINTS: PRO1271; HISDACETLASE.
 KW Hydrolase; Nuclear protein.
 SEQUENCE 513 AA; 57546 MW; C45387CF3A38906F CRC64;

Query Match 59.7%; Score 1517.5; DB 1; Length 513;
 Best local similarity 65.2%; Pred. No. 7.7e-118;
 Matches 283; Conservative 60; Mismatches 72; Indels 19; Gaps 6;
 QY 3 ADESGTSLPS-GPGRKRRRSYFEPTIGDYTGCGPRMRRIRMAHSLITVHILHRL 61
 Db 6 AGSGGNSLPSVGPGRKRRRSYFEPTIGDYTGCGPRMRRIRMAHSLITVHILHRL 65
 QY 62 EISRPDLASDIDGRFSPPEYVDPLASVSPESMGDPAAENLRFRNFGDCPVDFGLF 121
 Db 66 QYVNPANRERELRFRPAEETINFLRSYTPETQD--QILKRFNNGECVPYLDGLYSF 123
 QY 122 CRASAGSGTCAAVKLNRODADIAIMGGGLHAAKSEASGFCYVNDIVLGLLELLMKFKR 181
 Db 124 QYVYAGASVCGAVAFN--HGHDIAIMGGGLHAAKSEASGFCYVNDIVLGLLELLMKHER 182
 QY 182 VLYVIDVHGGDVEEAFYTTDRVMTVSFHKFGDPGTGTHIDVGAKEKKYALNVPRLN 241
 Db 183 VLYVIDIDIHGGDVEEAFYTTDRVMTVSFHKFGDPGTGTHIDVGAKEKKYALNVPRLD 242
 QY 242 DGMDESFRLSFLPRLIOKMEVYQPEAVNLOCGADSLSGRLGCFNLVSGHACDLRELSR 301
 Db 243 DGIDDESYGLHLLFPIKMKVMEIFRPGAVNLOCGADSLSGRLGCFNLVSGHACDLRELSR 302
 QY 302 SYNVPLNVLGEGYTTIRNVARCWCYETAFAVGYEPDNKLPYNEFEYFGPDYTLHVPSP 361
 Db 303 SFNVPLLLGLGGGYTTIRNVARCWCYETAFAVGYEPDNKLPYNEFEYFGPDYTLHVPSP 362
 QY 362 MENLTPKDMERIRNTLLEQSLGILHAPSVQFOHTPPVNV--LDEPEDMETRPKPRX 419
 Db 363 MENLTPKQDLDDIYS---KLKLRHAPSVFQERPVTEPREVEDQEDGDRMPDPS-- 415
 QY 420 SGTATYESDDDD 433
 Db 416 ----DSMDVVD 423
 RESULT 3
 HDAL_CHICK STANDARD; PRT; 480 AA.
 AC P56517;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HISTONE DEACETYLASE 1 (HD1).
 GN HDAC1 OR HDAC1A.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 NC NCHI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takami Y.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

OY 431 DDDKPLHGYSCRGATTDTRDSTGEDEMDNDPE 463
 DB 423 SDEEGEGRRKSSNFKKAKRYKTEDEKEDPEE 455
 RESULT 5
 HDAL_MOUSE STANDARD; PRT: 488 AA.
 ID HDAL_MOUSE
 AC P70288;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HISTONE DEACETYLASE 2 (HD2) (Y1) TRANSCRIPTION FACTOR BINDING
 DE PROTEIN).
 GN HDAC2 OR Y1BP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoma;
 RX MEDLINE=97075080; PubMed=8917507;
 RA Yang W.-M., Inouye C.J., Deng Y., Beatty D., Seto E.;
 RT "Transcriptional repression by Y1 is mediated by interaction with a
 RT mammalian homolog of the yeast global regulator Rpd3.*"
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY).
 CC -1- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
 CC AT LEAST WITH THE ZINC-FINGER TRANSCRIPTION FACTOR Y1.
 CC -1- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX WITH Y1.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U01758; AAC52889.1; -
 DR MGD; MGI:1097691; Hdac2.
 DR InterPro: IPR000286; His_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl1.
 DR PRINTS: PRO1270; HDASUPER.
 DR PRINTS: PRO1271; HISDACETLASE.
 KM Hydroxylase: Nuclear protein.
 FT DOMAIN 300 303 POLY-GLY.
 ST SEQUENCE 488 AA; 55302 MW; B9843D2A475157C CRC64;
 SQ
 Query Match 57.6%; Score 1455.5; DB 1; Length 488;
 Best Local Similarity 60.3%; Pred. No. 1.4e-113;
 Matches 270; Conservative 68; Mismatches 83; Indels 27; Gaps 6;
 OY 16 GRRKRVSYFEPTIGDYVYGCGHMKPRIRIMASHLIHYHLHRLRLEISPSLADASDIG 75
 DB 7 GCKKVCYTYGDIIGNYTYGCGHMKPRIRIMTNLLNLGLYKMKELTPHAKATAEMT 66
 OY 76 RPHSPRYVDPLASVSPESMCDPSAARNLRRNVGDEDCPVEDGDFDCFRASAGSIGAAVK 135
 DB 67 KYHSDIEIKFIRKIRPDNMSEYS--KOMORFNVEDCPVEDGDFEFCOLSTGSGVAAVAK 124
 OY 136 LNRDADAIIMWGGGILHAKKSEASGFCYVNDIVLGLLELLKMKRRLYLIDIVHDDGV 195
 DB 136 LNRDADAIIMWGGGILHAKKSEASGFCYVNDIVLGLLELLKMKRRLYLIDIVHDDGV 195

DB 125 LNRQOTDMAVNMAGLHIAKSSASGFCYVNDIVLAILLELKHORYLYIDIDIHGGCV 184
 OY 196 EEAFTYTRVMTVSFHKFGDFPPGTHIRDYGAENKRYALNPLNDGMDSEFSRSLFRP 255
 DB 185 EEAFTYTRVMTVSFHKFGDFPPGTHIRDYGAENKRYALNPLNDGMDSEFSRSLFRP 244
 OY 256 LIOKMEVYQPEAVNVIQCGADSLGDRIGCFNLISVGHADCLFRLSYVPLMVLAGEBY 315
 DB 245 LISKVEMVQPSAVNVLQCGADSLGDRIGCFNLISVGHADCLFRLSYVPLMVLAGEBY 304
 OY 316 TINNVARCMVEFAVAVGVPEDNKLPYNEFEYFGDYTLHVPSPMENLTPKMERIR 375
 DB 305 TINNVARCMVEFAVAVGVPEDNKLPYNEFEYFGDYTLHVPSPMENLTPKMERIR 364
 OY 376 NTLEQISGLIHAPSVQFHTPPVNRVLDPEDEMDETPRKPRXSGTATYESDSDDDKP 435
 DB 365 ORLFENLRMLPHAPGVQMAIPE----DAVHED-----SG----DEGDEPDKR 405
 OY 436 LHGYSCRGATTDTRDSTGEDEMDNDPE 463
 DB 406 L---SIR---ASDKRIACDEEFSDE 427
 RESULT 6
 HDAL_HUMAN STANDARD; PRT: 482 AA.
 ID HDAL_HUMAN
 AC Q13547; Q92534;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HISTONE DEACETYLASE 1 (HD1).
 GN HDAC1 OR RPD3L1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=96185499; PubMed=8602529;
 RA Taunton J., Hassig C.A., Schreiber S.L.;
 RT "A mammalian histone deacetylase related to the yeast transcriptional
 RT regulator Rpd3p.*"
 RL Science 272:408-411(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=96244606; PubMed=8646880;
 RA Furukawa Y., Kawakami T., Sudo K., Imazawa J., Matsumine A.,
 RA Akiyama T., Nakamura Y.;
 RT "Isolation and mapping of a human gene (RPD3L1) that is homologous to
 RT RPD3, a transcription factor in *Saccharomyces cerevisiae*.*"
 RL Cytogenet. Cell Genet. 73:130-133(1996).
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
 CC -1- SUBUNIT: FORMS A COMPLEX WITH RBA48 AND ALSO WITH THE HISTONE
 CC ACETYLTRANSFERASE PCAF AND THE ADAPTOR PROTEIN P300. ALSO FORMS
 CC A COMPLEX WITH SIN3 AND SAP18. INTERACTS WITH TGIF.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: UBICUITOUS, WITH HIGHER LEVELS IN HEART,
 CC PANCREAS AND TESTIS, AND LOWER LEVELS IN KIDNEY AND BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE HISTONE DEACETYLASE 1-2 (HD1) (RPD3 HOMOLOG).
 OS Xenopus laevis (African clawed frog).
 OC Amphibia: Batrachia: Anura: Mesobatrachia: Pipidae: Pipidae;
 OC Xenopodidae: Xenopus.
 NCBI_TaxID=8355;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RA Patterson D., Noll A.P.
 RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF020658; AAC60346.1;
 DR InterPro: IPR000286; Hist_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl_1.
 DR PRINTS: PRO1270; HDASUPER.
 DR PRINTS: PRO1271; HISTDECTELASE.
 KW Hydroxylase; Nuclear protein.
 FT DOMAIN 299 302 POLY-GLY.
 FT SEQUENCE 480 AA: 54892 MW: CA92DE34D3639E8 CRC64;
 Query Match 57.1%; Score 1453.5; DB 1; Length 480;
 Best Local Similarity 58.7%; Pred. No. 1.4e-112;
 Matches 263; Conservative 79; Mismatches 79; Indels 27; Gaps 6;
 QY 16 GKRKRVSYFEPTIGDYVYGOGHPMKPHRIRMAHSLIIHYHLHRLLEISRPDLASDIDG 75
 DB 6 GTKKKVVYDDGVGNYGOGHPMKPHRIRMAHSLIIHYHLHRLLEISRPDLASDIDG 75
 QY 76 RHSPHYVDPLASVPSMGPDSAAANLRRFNVGEDCPVFGCLPFCASAGSGIAGAVK 135
 DB 66 KYHSDDYIKFIRNIRPDNNSEYS--KOMORFNVGEDCPVFGCLPFCASAGSGIAGAVK 123
 QY 136 LNRQDADAIIMGGGLHAKKSEASGFCYVNDIVYGLIELLMKFRVLYTIDVHNGDV 195
 DB 124 LNKQOTDLSVNNSSGLHAKKSEASGFCYVNDIVYGLIELLMKFRVLYTIDVHNGDV 183
 QY 196 EBAFTTTRVMTVSFHAKGDFPFGTGHIRDVGAERKGYALANLPLNDGMDSEFSRFRP 255
 DB 184 EBAFTTTRVMTVSFHAKGDFPFGTGHIRDVGAERKGYALANLPLNDGMDSEFSRFRP 243
 QY 256 LIOKMEYVOPRAVYVLOGGASDLGRCFNLSYKGAOCLRFLRSVNVNVLVYGGG 315
 DB 244 VMTKMEHOPRAVYVLOGGASDLGRCFNLSYKGAOCLRFLRSVNVNVLVYGGG 303
 QY 316 TIRNARCMCEYTAAYVCEPDNKLPRNEFEYFGPDYTLHVDSPMNLTPKMERIK 375
 DB 304 TIRNARCMCEYTAAYVCEPDNKLPRNEFEYFGPDYTLHVDSPMNLTPKMERIK 363
 QY 376 NTLLFQLSGLIHAPSVQFOHPPVNRVLDEPEDMETRPPKRXSGTATYSSDDDDKP 435
 DB 364 QRLFENLHRLPHAPGVQMAIPE-----DSVIDD-----SG-----EDEDDPDKR 404
 QY 436 LHGSCRGATITDRDSTGEDEMDNDNP 463

DB 405 I---SIR---SSDKRIACDEEFSDE 426
 RESULT 9
 HDAC_DROME STANDARD: PRT: 520 AA.
 ID HDAC_DROME
 AC Q94517; 017429;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PROBABLE HISTONE DEACETYLASE (HD) (DRPD3).
 GN RPD3 OR HDAC1.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;
 OC Ephydroidea: Drosophilidae: Drosophila.
 NCBI_TaxID=7227;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RA Tissue-Embryo;
 RX MEDLINE=97113416; Pubmed=8955276;
 RA de Rubertis F., Kadosh D., Henchoz S., Pauli D., Reuter G., Struhl K.,
 RA Spleiter P.;
 RT The histone deacetylase RPD3 counteracts genomic silencing in
 RT Drosophila and yeast.
 RL Nature 384:589-591(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johnson C.A., White D., O'Neill L.P., Turner B.M.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
 CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
 CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
 CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
 CC SIMILARITY). IN DROSOPHILA, IT IS INVOLVED IN POSITION-EFFECT
 CC VARIATION (PEV) AND IS DIRECTLY INVOLVED IN COUNTERACTING
 CC GENOME SILENCING.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y09258; CAAT0455.1;
 DR EMBL: AF026949; AAC23917.1;
 DR FlyBase: FBgn0015805; Rpd3.
 DR InterPro: IPR000286; Hist_deacetylase.
 DR Pfam: PF00850; Hist_deacetyl_1.
 DR PRINTS: PRO1270; HDASUPER.
 DR PRINTS: PRO1271; HISTDECTELASE.
 KW Hydroxylase; Nuclear protein.
 FT CONFLICT 50 51 DI -> ELY (IN REF. 2).
 FT CONFLICT 66 66 C -> S (IN REF. 2).
 FT CONFLICT 96 96 D -> N (IN REF. 2).
 FT CONFLICT 105 105 E -> D (IN REF. 2).
 FT CONFLICT 295 295 V -> VV (IN REF. 2).
 FT CONFLICT 370 370 L -> V (IN REF. 2).
 FT CONFLICT 506 506 S -> T (IN REF. 2).
 FT CONFLICT 520 520 S -> T (IN REF. 2).
 FT SEQUENCE 520 AA: 58183 MW: D02EA3DBD3C64688 CRC64;
 Query Match 57.1%; Score 1453.5; DB 1; Length 520;
 Best Local Similarity 57.9%; Pred. No. 1.5e-112;
 Matches 270; Conservative 77; Mismatches 90; Indels 29; Gaps 6;
 QY 18 KRVSYFEPTIGDYVYGOGHPMKPHRIRMAHSLIIHYHLHRLLEISRPDLASDIDG 77

```

Db 6 KKRVCYDDSDIGNYYGQGHMKPRLIRMTNLLNGLYRKMDI--RPHKATDDEMFKF 64
OY HSPVVDPLASVSPSPDPSAARNLRPNVNGEDCPVDFGDFPCRASAGSIGAAVYLN 137
Db 65 HCDYVRLRLSRIRPDNMEYN--KOMQFNVGDCPPVDFGLYFCQLSAGSVAAAYKLN 122
OY 138 RODADIAIMWGGLHAKKSEASGFCYVNDIYLIGLELLMKFRVLYIDIVHGDGVEE 197
Db 123 KQASEICIMWGGLHAKKSEASGFCYVNDIYLIGLELLMKFRVLYIDIVHGDGVEE 182
OY 198 AFYTTDRMTVTSFHKFGDFPGTGHIRVDGAERKGYVALNPLNDGMDDESPRSLRPLI 257
Db 183 AFYTTDRMTVTSFHKFGDFPGTGHIRVDGAERKGYVALNPLNDGMDDESPRSLRPLI 242
OY 258 QKRVNVEYQPEAVYVLOCAGDSLGCDFNLGVKSHADCLRFLRSYVPLVWLGEGYTI 317
Db 243 SKVNETFPPAAVYVLOCAGDSLGCDFNLGVKSHADCLRFLRSYVPLVWLGEGYTI 302
OY 318 RNVAACWCYETAVALVGEVDPDKLPYNEVEYFGPDYTLHVPSPMENLTPKDMERIRNT 377
Db 303 RKVSHCWYETETVALVGEVDPDKLPYNEVEYFGPDYTLHVPSPMENLTPKDMERIRNT 362
OY 378 LLEQLSLGILHAPSVOFQHTP--PYNRVLDE-----PEDMETRPRKRXMSGTATY 425
Db 363 LFEINLRMLPHAPGVQIQAIPEDAIINDESDDEKDKYDKDRLQSDSKRIYPE-----NE 416
OY 426 ESDSDDDKPLHGYSCRGATTDRTDSDDEMDNDNPEPDVNPSS 471
Db 417 YSDESEDE-----GGRDNRSYKQGRKRPRLDKDTNKNASS 454

RESULT 10
HDAL_HUMAN STANDARD: PRT; 488 AA.
AC Q92769;
DT 15-JUL-1998 (Rel. 36, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE HISTONE DEACETYLASE 2 (HD2).
GN HDAC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-Breast;
RA MEDLINE-97075080; PubMed-8917507;
RA Yang W.-M., Inouye C.J., Zeng Y., Beards D., Seto E.;
RT "Transcriptional repression by YY1 is mediated by interaction with a
RT mammalian homolog of the yeast global regulator RPD3."
RL Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -1- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC WITH MAD, SIN3, YY1 AND N-COR.
CC -1- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; LOWER LEVEL IN BRAIN AND
CC LUNG.
CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY, HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/)

```

```

CC or send an email to license@isb.ch.
CC -----
CC EMBL: U31814; AAC50814.1; -
CC MIM: 605164; -
CC DR Interpro: IPR000286; His_deacetylase.
CC DR Pfam: PF00850; His_deacetylase.
CC DR PRINTS: PR01270; HDASUPER.
CC DR PRINTS: PR01271; HISDACETLASE.
CC FT Hydroxylase; Nuclear protein.
CC FT DOMAIN 300 303 POLY-GLY.
CC SQ SPOUNCE 488 AA; 55325 MM; 3AB554ADC7734B70 CRC64;

Query Match 56.8%; Score 1445.5; DB 1; Length 488;
Best Local Similarity 59.6%; Pred. No. 6.4e-112;
Matches 267; Conservative 68; Mismatches 86; Indels 27; Gaps 6;

OY 16 GKRKRVSTFTEPTIGDYTYGQGHMKPRLIRMAHSLIHYHLHRLRISPLADSDIG 75
Db 7 GQKRVYVYDGDIGNYYGQGHMKPRLIRMTNLLNGLYRKMEIYRPHKATAEEMT 66
OY 76 RPHSPEYVDPLASVSPSPDPSAARNLRPNVNGEDCPVDFGDFPCRASAGSIGAAVY 135
Db 67 KYHSDYIKFLRSIRPDNMEYS--KOMHIFNVGDCPPVDFGLYFCQLSAGSVAAAYK 124
OY 136 LNRDADIAIMWGGLHAKKSEASGFCYVNDIYLIGLELLMKFRVLYIDIVHGDGCV 195
Db 125 LNRQOTDMAVWAGGLHAKKSEASGFCYVNDIYLALLELLKYHQRVLYIDIVHGDGCV 184
OY 196 EEARVTTDRVMTVTSFHKFGDFPGTGHIRVDGAERKGYVALNPLNDGMDDESPRSLR 255
Db 185 EEARVTTDRVMTVTSFHKFGDFPGTGHIRVDGAERKGYVALNPLNDGMDDESPRSLR 244
OY 256 LQKRVNVEYQPEAVYVLOCAGDSLGCDFNLGVKSHADCLRFLRSYVPLVWLGEGY 315
Db 245 LITSVMEYQPSAAVYVLOCAGDSLGCDFNLGVKSHADCLRFLRSYVPLVWLGEGY 304
OY 316 TIRNVAACWCYETAVALVGEVDPDKLPYNEVEYFGPDYTLHVPSPMENLTPKDMERIR 375
Db 305 TIRNVAACWCYETAVALDCEIPELNPYNDYFEGPDKLHISPSNNTNONTPEYMKIK 364
OY 376 NTLLEQLSLGILHAPSVOFQHTP--PYNRVLDEPEDMETRPRKRXMSGTATYESD 435
Db 365 QRLFEINLRMLPHAPGVQIQAIPEDAIINDESDDEKDKYDKDRLQSDSKRIYPE-----SG-----DEGDEPD 405
OY 436 LHGYSCRGATTDRTDSDDEMDNDNPE 463
Db 406 L---SIR---ASDKRIACDEEFSDEDE 427

RESULT 11
HDAL_STRPU STANDARD: PRT; 576 AA.
AC P56518;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 1 (HD1).
GN HDAC1.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID:7668;
RN 11
RP SEQUENCE FROM N.A.
RA Nemer M.;
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:45 ; Search time 25.88 Seconds

(Without alignments)
1386.330 Million cell updates/sec

Title: US-09-645-337-4

Perfect score: 25.44

Sequence: 1 MEADSGSLSPGPDGPKRRR.....TGEDEMDDDNPPPPVNPSS 471

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1560.5	61.3	501	2 T05640	histone deacetylase
2	1517.5	59.7	513	2 T05640	probable histone d
3	1456.5	57.3	480	1 S60381	RPD3 protein homol
4	1359.5	53.4	461	2 T20163	hypothetical prote
5	1348.5	52.2	433	1 S22284	transcription regu
6	1305.5	51.3	465	2 T23963	hypothetical prote
7	1292.5	50.8	424	2 JC7102	histone deacetylase
8	1285.5	50.5	424	2 JC5834	histone deacetylase
9	1275	50.1	405	2 T40300	histone deacetylase
10	1207	47.4	419	2 T47443	hypothetical prote
11	1143.5	44.9	434	2 T11643	histone deacetylase
12	1041.5	40.9	507	2 T19067	hypothetical prote
13	1003.5	39.4	452	2 S64211	probable transcript
14	457.5	18.0	375	2 D70388	acetylpolymyosin
15	436	17.1	470	2 S54089	hypothetical prote
16	434.5	17.1	367	1 S39643	acetylpolymyosin
17	424.5	16.7	367	2 T36278	acetylpolymyosin
18	410.5	16.1	359	1 B69266	acetylpolymyosin
19	342	13.4	359	1 B69266	acetylpolymyosin
20	319.5	12.6	380	2 D83174	hypothetical prote
21	309.5	12.2	706	2 S62933	hypothetical prote
22	305	12.0	310	1 A70481	hypothetical prote
23	299.5	11.8	158	2 T48929	hypothetical prote
24	269.5	10.6	425	2 G86217	protein T2767.14
25	269.5	10.6	796	2 T32425	hypothetical prote
26	268	10.5	338	2 H84173	acetylpolymyosin
27	265	10.4	878	2 T17245	hypothetical prote
28	264	10.3	687	2 T43797	probable histone d
29	261.5	10.3	343	1 G64366	acetylpolymyosin

30	255.5	10.0	306	2 D82126	histone deacetylase
31	251.5	9.9	335	1 H71071	hypothetical prote
32	251.5	9.9	1095	2 T13964	probable histone d
33	245.5	9.7	331	1 C69026	acetylpolymyosin
34	243.5	9.6	385	2 A81926	hypothetical prote
35	239.5	9.4	142	2 T47441	hypothetical prote
36	239.5	9.4	1063	2 T46284	hypothetical prote
37	235.5	9.3	369	2 F81178	histone deacetylase
38	232.5	9.1	310	2 B47050	histone deacetylase
39	232	9.1	304	1 S74557	gluc 3'-region hyp
40	231.5	9.1	334	2 B75095	acetylpolymyosin
41	229	9.0	517	2 T27101	probable histone d
42	215.5	8.5	301	2 H75470	hypothetical prote
43	215.5	8.5	359	2 T05998	histone deacetylase
44	203	8.0	346	2 H83469	hypothetical prote
45	203	8.0	782	2 T22134	acetylpolymyosin

ALIGNMENTS

RESULT	1	
T05640	histone deacetylase (EC 3.5.1.-) F20D10.250 - Arabidopsis thaliana	
N:Alternate names:	protein F20D10.250	
C:Species:	Arabidopsis thaliana (mouse-ear cress)	
C:Date:	23-Apr-1999	#sequence_revision 23-Apr-1999 #text_change 12-Nov-1999
C:Accession:	T05640	
R:Evans, M.; Medler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999		
A:Reference number:	215420	
A:Accession:	T05640	
A:Molecule type:	DNA	
A:Residues:	1-501 <BEV>	
A:Cross-references:	EMBL:AL035538; GSPDB:GN00062; ATSP:F20D10.250	
A:Experimental source:	cultivar Columbia; BAC clone F20D10	
C:Keywords:	hydrolase	
F:21-320/Domain:	RPD3/acuc homology <RAH>	
Query Match	61.3%; Score 1560.5; DB 2; Length 501;	
Best Local Similarity	60.7%; Pred. No. 2.8e-116;	
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;		
QY	4 DEGSISLPSGPDGPKRRRVSFEPTGDDYVGGGHMKPRIRMAISLIHHHLRLRI 63	
DB	2 DTGNSLASGPDGPKRRVCFYDPEVGNYYGGGHMKPRIRIMTIALAHVGLDHMOV 61	
QY	64 SRPSLADASDGRFHSPEYVDFLASVSPESMGDPASARINRPNVEDCVFPGFLDFPR 123	
DB	62 LKFPFADRLDLCRPHADYVSLRSLITPEQQD--QIRQLRRNVEDCVFPGFLSFCQ 119	
QY	124 ASAGSGTGAAYVKNRQDADIAINMGGLHAKKSEASGFCYVNDIVLGIIELEKFRVYL 183	
DB	120 TVAGSGVSGSVKLNHGLCDIAINMGGLHAKKSEASGFCYVNDIVLGIIELEKHERVYL 179	
QY	184 YIIDIHNGGVEAEAYTTDRVMTVSFHKRGDFPFGHTRDVGAKRGYVYALNVLNNG 243	
DB	180 YVIDIHNGGVEAEAFATDRVMTVSFHKRGDFPFGHTRDVGAKRGYVYALNVLNNG 239	
QY	244 MDSEFSKSLRPLIKVMEVYQPEAVVLQCADSLSGDRICFNLVSKGHADLRLRSY 303	
DB	240 IDDESHTLTKPKIMKGMELFRGAVVILQCADSLSGDRICFNLVSKGHADLRLRSY 299	
QY	304 NVPLMTVIGGGYTRIRNARCWCETAVANGVBPDKLPIYNEVFEYGPVYTLHVDSPME 363	
DB	300 NVPLMTVIGGGYTRIRNARCWCETAVANGVBPDKLPIYNEVFEYGPVYTLHVDSPME 359	

A; Introns: 58/3; 98/1; 387/1
C: Superfamily: RPD3 protein; RPD3/acuc homology

Query Match	53.48	Score 1359.5	DB 2	Length 461
Best Local Similarity	57.38	Pred. No. 2.4e	100	
Matches 262	Conservative	65	Mismatches 109	Indels 21
				Gaps 5

```

QY 12 SGP DGRKRRSYSEYPTIGDYGGGCHPMKPHRRMASHLIIHYHLRRLEISPSLA 69
Db 4 NGLPMHEKRRVAYIYDSDNIGYITGGCHVMKPHRRMTHLHLVLTGLRLEIFRPPA 63
QY 70 DASDIGREHSEYVDPLASVSPESMGDPSAARNLRNFNVGEDCPYFDGLFDCRASAGS 129
Db 64 SEEDMTFRHSEDEYMTFLKSANPND- -KSFNMQLKFENVGEDCPLEFDGLYERCOLSSGS 121
QY 130 IGAAYKLRRQDADIAIINMGCGLHHAKKSSAFSCYVNDIYVLGILELLKMFKRYLIYDIDY 189
Db 122 LAAATKLKKOKVADIAIINMGGLHHAKKSEASGEFCYTNIDVLGILELLKHKRRLYVYDIDY 181
QY 190 HHGCGVEAEFYTTDRMYTVSFKHFGDFPGTGHIRDVGAEKQKYVALPLNDGMDSEF 249
Db 182 HHGCGVEAEFYTTDRMYTVSFKHFGDFPGTGLNDIGAKGKRLSVNPLRDGIDYVSTY 241
QY 250 RSLFRPLIÖKWEVYÖPEAVVYLOGGADSLSGDRLCGFNLVYKRGHADCJLFLRSYVNPVLY 309
Db 242 QSIFKPIMTKWEKRPDCAVVYLOGGADSLNGDRLGFNLTLKGHGECARFFRSYVNPVLY 301
QY 310 LGGEGYTIIRNARCYCEAVAVGVEPDKLPLYNEXFEFGDYTLHYHPSPRENINTFK 369
Db 302 VGGGYTIRNARCYTEIETSIANDKEVPNELPYNDIEFEFGNYRIHIISSNANANSSSD 361
QY 370 DMERIRNLTILQOLSLIHAPSVYOFHTPPVNRVLDPEDEDMETRPKPRXMSGATAYESDS 429
Db 362 MLAKIQTDVIANTLEQTLFVPSVQMRPI--EDALSLAND-----SLADQ 405
QY 430 DDDDKPLHGYSRGGATTDRO-STGEDEMDONPREP 465
Db 406 ANPDKRLPPQITDGMIDDGDFYDGEREGDOKRNEED 442

```

RESULT 5
S22284
transcription regulator RPD3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: Protein N0305; Protein YNL330c
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence,revision 10-Sep-1999 #text_change 23-Mar-2001
A:Accession: S22284; S51336; S51285; S55859; S57393; S63313; S63311
R:Vidal, M.; Gaber, R.F.
Mol. Cell. Biol. 11, 6317-6327, 1991
A:Title: RPD3 encodes a second factor required to achieve maximum positive and negative
A:Reference number: S22284; MUID:92049362
A:Accession: S22284
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-433 <VID>
A:Cross-references: GR:S66438; NID:q238961; PIDN:AAB20328.1; PID:q238962
R:van Dyck, L.; Pascual-huiri, A.; Goffeau, A.
submitted to the EMBL Data Library, December 1994
A:Description: A 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes
located gene for a putative aryl-alcohol dehydrogenase.
A:Reference number: S51334
A:Accession: S51336
A:Molecule type: DNA
A:Residues: 1-433 <VAN>
A:Cross-references: EMBL:X81226; NID:g642335; PIDN:CAA58228.1; PID:g642338
R:Nicaud, J.J.
submitted to the EMBL Data Library, January 1995
A:Description: Sequence analysis of a 13.9 Kb fragment of yeast chromosome XIV identifying
A:Reference number: S51285
A:Accession: S51285
A:Molecule type: DNA
A:Residues: 1-201 <NIC>

A:Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86368.1; PID:g854536
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Galliardin, C.

	Accession	Reference number	Assignment Of yeast chromosome	Identified
A:Status:	nucleic acid sequence not shown;	translation not shown		
A:Accession:	S5859	MUID:95373280		
A:Reference number:	S5859			
A:Assignment:	Yeast chromosome XIV			

A:Residues: 1-201 <MAM>
A:Cross-references: EMBL:D46259; NID:9633655; PIDDN:CAA6368.1; PID:9854536
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1995
R:van Dyck, L.; Pascal-Ahuir, A.; Punelle, B.; Goffeau, A.
Yeast 11, 987-991, 1995
A>Title: An 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes as
d gene for a putative aryl-alcohol dehydrogenase.
A:Reference number: S57391; MUID:96021610
A:Accession: S57393
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-433 <VAM>
A:Cross-references: EMBL:X83226; NID:9642335; PIDDN:CAAS8228.1; PID:9642338
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, December 1994
R:Coste, F.; Jomiaux, J.L.; Goffeau, A.; Punelle, B.; Van Dyck, L.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S63151
A:Accession: S63153
A:Molecule type: DNA
A:Residues: 1-433 <COS>
A:Cross-references: EMBL:Z71606; NID:q1302450; PIDDN:CAAG6263.1; PID:q1302451; GSPDB:G:
A:Experimental source: Strain S288C
R:Matahi, M.; Nicaud, J.M.; Levesque, H.; Galliard, C.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S63287
A:Accession: S63311
A:Molecule type: DNA
A:Residues: 1-201 <MAW>
A:Cross-references: EMBL:Z71606; GSPDB:GN00014; MIPS:YNL330C
A:Experimental source: Strain S288C
A:Genetics:
A:Gene: SGD:RPD3; SD12; SDS6; MIPS:YNL330C
A:Cross-references: SGD:S0005274; MIPS:YNL330C
A:Map position: 14L
C:Function:
A:Description: transcription regulation
C:Superfamily: RPD3 protein; RPD3/acuc homology
C:Keywords: nucleus; transcription regulation
F:23-322/Domain: RPD3/acuc homology <RAH1>

	Query Match	Best Local Similarity	Matches
OY	18 KRRVSEFEPTGGDYGYGOGHPMKPHRIIMASHLIHYHLIRLEISPSLADASDIGRF 77	52.2%; Score 1328.5; DB 1; Length 433;	Conservative 74; Mismatches 94; Indels 5; Gaps 2;
Ddb	18 KRRVAFFVDADVGNATAYAGAHMPKPHRIIMASHLIIMNGLYKKMEIYAKPATPKEMQOF 77	57.9%; Pred. NO. 6.4e-98;	
OY	78 HSPEYVDVLASVSPESMDDPSAARNLRFPNVGEDCPVPDGLPFCFRASAGSGTGAAYLN 137		
Ddb	78 HTDEYDFLSRYTPDNLT--EMPRRESVKFNFGDDCPVPFDGIYEYSISGGSGMECAARLN 135		
OY	138 RODADIAINMGSGGLHAHAKKSEASGFCYNDIVYLGLELLLMKFKRLVLDIDVHHDDGYEE 197		
Ddb	136 GKCKVAANVAGGLHNAKKSSEASGFCYNDIVYLGLELLLRHPRLVLDIDVHHDDGYEE 195		
OY	198 AFYTDRMTVTSFHFKGGEFFPGTGHIRDVGAERKUYVALNPDLNMGMDSEFSRLFLPL 257		
Ddb	196 AFYTDRMTGCSFHKHGGEFFPGTGELROIYGACAKNYAVNNPDLRGCIDDATYRSVFERY 255		
OY	258 OKWMEVYPAAVAVLCOGDSTLSGDRLGCFNLGVKSHADCLRFLRSYVNPVLMLVGEGYTI 317		
Ddb	256 KIKMWYPPSPAAVAVLCOGDSTLSGDRLGCFNLSEGHANCYVNVKSFGLPMWVVGGSYTM 315		

[illegible]

RESULT 13

Probable transcription regulator YGL194c - Yeast (*Saccharomyces cerevisiae*)
N.Alternate names: hypothetical protein G1330
C.Species: *Saccharomyces cerevisiae*
C.Date: 17-May-1996 #sequence.revision 17-May-1996 #text.change 20-Jun-2000
C.Accession: S64211; S62051
R.Bruschi, C.V.; Coglianina, M.; Bertani, I.; Klim, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A.Reference number: S64183
A.Accession: S64211
A.Molecule type: DNA
A.Residues: 1-452 <BRU>
A.Cross-references: EMBL:E272716; NID:g1322818; PID:CAA96906.1; PID:g1322819; MIPS:YGL194c
A.Experimental source: strain S286C
R.Coglianina, M.; Delneri, D.; Zaccaria, P.; Klim, R.; Bertani, I.; Bruschi, C.V.
submitted to the EMBL Data Library, September 1995
A.Description: A 6.7 Kb fragment from chromosome VII of *Saccharomyces cerevisiae* contain
A.Reference number: S62051
A.Accession: S62051
A.Molecule type: DNA
A.Residues: 1-351, 'NSGRHS', 358, 'RIIHPT', 365-452 <COG>
A.Cross-references: EMBL:X91837; NID:g1177627; PID:CAA62950.1; PID:g1177634
A.Experimental source: strain FY1679
C.Genetics:
A.Gene: SGD:HOS2
A.Cross-references: SGD:S0003162; MIPS:YGL194c
A.Map position: 7L
C.Superfamily: RPD3 protein, RPD3/acuc homology
#30-331/Domains: RPD3/acuc homology <RAHL>

Query Match	39.4%	Score 1003.5	DB 2	Length 452
Best Local Similarity	44.0%	Fred. No. 4.8e-72		
Matches 198	Conservative 86	Mismatches 127	Indels 39	Gaps 9
QY	20	RVSYYEPTIGDYYGGGHPMKPHRIEMAHSLIHYYLHRRLEISRPSLADASDIGRPHS	79	
		: : : : : : : : :		
Db	27	RVSYHFNKSVSHYHGVAHHPMKPFLMLTLDLHVSYGHLKIMDLVETRASATDELLQFHS	86	
		: : : : : : : : :		
QY	80	PEYVDFLASPESMGDSAAARLIRFVNGEDCPAFEGDIFEDCRASAGSICAAKLNQ	139	
		: : : : : : : : :		
Db	87	EDYVNFELSKVSPEN-ANKLPKGTLENFMGIDCCPFLQNYDITLYTGTASLDATKRLINN	145	
		: : : : : : : : :		
QY	140	DADIAINNGGGLHNAKKEASEGFCYVNDIVLGIETLLKMKFVLYIDIVHGGIDVEEAF	199	
		: : : : : : : : :		
Db	146	QSDIAINNNGGLHNAKKNKSPSGFCYVNDIVLSTMLLRHPRLIYIDIDLHHGGDQGEAF	205	
		: : : : : : : : :		
QY	200	YTTDFKVTVSFHKF-GDFEPFGTGHIRDYGAEEKKYYALVPLNDMDDESEFSLRPLQ	258	
		: : : : : : : : :		
Db	206	YTTDFVFLTSPFKYNGEFPFGTGDLTLEIGCDGKRFALNVPLEDDIDDDSYINLFKSIYD	265	
		: : : : : : : : :		
QY	259	KVMEYVEQPAVYVLOGGADSLSGDRGECNLSYKGHADCLRFLRSYNVPLVMLVGGEGYIIR	318	
		: : : : : : : : :		
Db	266	PLIMFKEPFLTVIOGGGADSLGHDRLGCFNLKANGCECKAFKSNKGGLPVLVYGGGGYTPR	325	
		: : : : : : : : :		
QY	319	NVAKWCETIATAVAVGEVDNKLPIVNEYF-EYFGDYTLAHVDSPM----ENLNTPKOME	372	
		: : : : : : : : :		
Db	326	NVSRLMTYETGILNDVLLPEDIPEDIPEDSPGSPOLY----PMIDLVEKKNKKLLE	381	
		: : : : : : : : :		
QY	373	RIRNTLEQLSLIHAPSVOGHTPPVNRVLDEPEDOMETRKPKXWGSTATAYESDSDD	432	
		: : : : : : : : :		

```

Dd      382 DIRI RCL ENI NY L G A D S V R M-----D A C I P-----T Q D I S A L T E E 419
QY      433 DK P L H G Y S C R G A T T D R D S T G E - D E M O D D N 461
         ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Dd      420 DK I T Q E M N-----E E F E A D S S N R L E E M E K E N 445

```

RESULT

acetoIn utilization protein - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
 C:Accession: D70388
 R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 N:ature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: D70388
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-375 <GB>
 A:Cross-references: AB:AE000719; MID:g2983517; PIDN:AAC07100.1; PID:g2983524; GB:AE000719
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: acuC1
 C:Superfamily: acetylpolylamine aminohydrolase: RPD3/acuC homology
 I:6-306/Domain: RPD3/acuC homology <RAH1>

Query Match	18.08;	Score 457.5;	DB 2;	Length 375;
Best Local Similarity	34.08;	Pred. No. 9.7e-29;		
Matches 127; Conservative	65;	Mismatches 140;	Indels 41;	Gaps 14

QY 30 CDYVGGGHHMKHRIIMAHSLIHHLRKEISRPSLADSDIGRHFSEYVDLASY 89
DB 13 GKTYFKNHKLPRSLRLFRDAMNLIDEKELKSRPKRKEELLCFHEDYVNTIME- 71
QY 90 SPBSMGDPSAARNLRRPNV-GEBCPYFDGLDFPCASAGSGTGAAYKLNRODADIATNMG 148
DB 72 AEROCQVFKAR--EKTNTGTGPNPSTAMFGSSLAGSGTVOALIEELK--GNVAANPA 127

Db 128 GGMHAFKSRANGFCINDPAVCI EYLKKGFIRLLYIDDAHHCQGVGEAFYDTDOYFV 187

266 PEAVVLQCGADSLSGDRIGCFNLGVKHAQDLREFRSVNPVAVY-----LGGEGYITRN 319
Db 248 PEVYLLQIGTLPDLIEDLYSKPNLS-----NVALKAFNIYREVFEFGESVYILGGGGYHPYA 301

```

      :| | : | : :| | : | | | | |
Db 302 LARAWTLWCELSGREYPEKLNKAKELIKSIDEFFDDE---VDRSYM--LETLKDPW 355

```

```

      : : : |
Db 356 RGGEVRKEVDTL 368

```

RESULT 15

hypothetical protein YP0686 - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein YP499.23c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
 R:Accession: S54089
 R:Badcock, K.: Churcher, C.M.
 submitted to the EMBL Data Library, May 1995

A:Reference number: S54059
 A:Accession: S54089
 A:Molecule type: DNA
 A:Residues: 1-470 <BAD>
 A:Cross-references: EMBL:249219; NID:9805025; PID:9805048; MIPS:YPR068C
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:HOS1
 A:Cross-references: SGD:S0006272; MIPS:YPR068C
 A:Map position: 16R
 C:Superfamily: RPD3/acuc homology
 F:9-383/Domain: RPD3/acuc homology <RAH1>

Query Match 17.1%; Score 436; DB 2; Length 470;
 Best local similarity 28.0%; Pred. No. 6.8e-27;
 Matches 104; Conservative 67; Mismatches 116; Indels 84; Gaps 10;

```

QY 39 PMKPH-RIRMAHSLIHHLHRL-EISRPLADASDTCRFHSPRYVDFLAS-----VSP 91
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 20 PCNNHOKSOLTYSLINAVDLQHFEDEVLTFFPYARKDLEFHSKSYIDYLLNGRFNKMA 79
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 92 ESMDDP-----SAARNLRRF----- 106
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 80 QDVNPNVYESKSELSELADNNKIDYNPSODLQRFTRNLNLYNLNHSQALENNMDC 139
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 107 -----NVEDCPVFDGLFPCRASAGSIGAAVKLRQDA 141
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 140 INNSEVPTNDKPTQYILNSETKQYNLEGDPIFSYLPWCQVITGATLNLDLHLSPTER 199
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 142 DIAIMGGGLHHAHAKSEASGFCYVNDIVLGILELK-MEKRYLYIDIDVHHGDCYEAFY 200
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 200 LIGIMMDGGRHHAFAKQASGFCYINDVYLLIQLRKAKLNTIYVDFDLHGDGVEKAFQ 259
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 201 TTDRTVMTVSFHKF-GDFPPTGTHIRDVGAEEKKYALNVPINDGMDDESFRSLFRPLQK 259
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 260 YSKQIOTISVHLIEPGFFPGTGSLSDSRKDN---VNIPLKHGCDNYLELIASKIVNP 316
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 260 VMEVYQPEAVYLQCGADSLSGDRLGCFNLSVKGHAD-CLREFLRSY-NVPLMVLGGEGYTI 317
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 317 LIEHHEPEALLIECGGDDGLDGRFNEMQUTIRGLSRILIIIMKSYRAHIFLLGGGGYND 376
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 318 RNVARCWCYET 328
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 377 LLMSRFTYLT 387
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
  
```

Search completed: April 26, 2002, 17:52:47
 Job time: 110 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:51:24 ; Search time 21.38 seconds

(without alignments)
495.746 Million cell updates/sec

Title: US-09-645-337-4

Perfect score: 2544

Sequence: 1 MEADSGISLSPGDPGRKR.....TGEDEMDNDPEPVNPSS 471

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTOS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	1734	68.2	458	4	US-09-282-305-2
2	1551.5	61.0	517	4	US-09-282-305-8
3	1520	59.7	439	4	US-09-282-305-6
4	1448.5	56.9	482	1	US-08-528-255A-1
5	1448.5	56.9	482	1	US-08-717-365-1
6	1280.5	50.3	351	4	US-09-282-305-4
7	1249.5	49.1	432	4	US-09-282-305-10
8	247.5	9.7	335	4	US-09-446-504-80
9	104.5	4.0	6095	4	US-09-144-085-2
10	104.5	4.0	4150	4	US-09-428-517-2
11	93	3.7	593	1	US-08-296-362-2
12	91	3.6	396	4	US-08-867-611-2
13	91	3.6	396	4	PCT-US92-06965A-7
14	88	3.5	820	1	US-08-291-896-2
15	88	3.5	820	1	US-08-485-278-2
16	87.5	3.4	473	4	US-09-286-691-15
17	87.5	3.4	473	4	US-08-687-147-15
18	87.5	3.4	1135	2	US-08-469-537A-97
19	86	3.4	4928	4	US-09-036-987A-5
20	86	3.4	4928	4	US-09-370-700-5
21	85.5	3.4	767	1	US-08-446-777-6
22	85	3.3	438	3	US-08-486-099-105
23	85	3.3	438	3	US-08-360-107A-115
24	85	3.3	438	3	US-08-484-223B-105
25	85	3.3	438	3	US-08-919-597-105
26	85	3.3	438	3	US-08-475-668A-105
27	85	3.3	438	3	US-08-485-551A-105

28	85	3.3	438	3	US-08-471-913A-105	Sequence 105, App
29	85	3.3	438	4	US-08-485-264A-105	Sequence 105, App
30	85	3.3	550	1	US-08-279-700-18	Sequence 18, App1
31	85	3.3	550	1	US-08-279-700-20	Sequence 20, App1
32	85	3.3	550	1	US-08-279-700-22	Sequence 22, App1
33	85	3.3	550	4	US-09-230-944-20	Sequence 20, App1
34	84.5	3.3	767	1	US-08-446-777-8	Sequence 8, App1
35	84	3.3	541	4	US-08-867-611-8	Sequence 8, App1
36	84	3.3	541	5	PCT-US92-06965A-13	Sequence 13, App1
37	83.5	3.3	435	1	US-08-484-105-12	Sequence 12, App1
38	83.5	3.3	435	1	US-08-484-105-11	Sequence 11, App1
39	83	3.3	417	4	US-08-867-611-20	Sequence 20, App1
40	83	3.3	417	5	PCT-US92-06965A-25	Sequence 25, App1
41	83	3.3	550	1	US-08-348-891A-5	Sequence 5, App1
42	83	3.3	550	2	US-08-905-817-5	Sequence 5, App1
43	83	3.3	933	3	US-08-293-728-2	Sequence 2, App1
44	83	3.3	933	4	US-09-421-868-2	Sequence 2, App1
45	83	3.3	3567	2	US-07-642-734C-4	Sequence 4, App1

ALIGNMENTS

RESULT 1									
US-09-282-305-2									
Sequence 2, Application US/09282305									
Patent No. 6287843									
GENERAL INFORMATION:									
APPLICANT: Baldwin, Donald A.									
APPLICANT: Briggs, Steven P.									
APPLICANT: Crane, Virginia C.									
TITLE OR INVENTION: Maize Histone Deacetylases And Their Uses									
FILE REFERENCE: 5718-44,									
CURRENT APPLICATION NUMBER: US/09/282,305									
CURRENT FILING DATE: 1999-03-31									
PRIOR APPLICATION NUMBER: 60/080,563									
PRIOR FILING DATE: 1998-04-03									
NUMBER OF SEQ ID NOS: 18									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 2									
LENGTH: 458									
TYPE: PRT									
ORGANISM: Zea mays									
US-09-282-305-2									
Query Match									
Best Local Similarity 68.2%; Score 1734; DB 4; Length 458;									
Matches 328; Conservative 50; Mismatches 73; Indels 22; Gaps 7;									
QY	1	MEADSGISLPS---	GPGRKRKRVSYFYEPTTGGDYGGGHPMKPRIRIMASHLIIHYHL	57					
DB	1	MAASGASLSPGAGDEAHRRVSYFEPISGDIYGGGHPMKPRIRIMASHLVVHYXL	60						
QY	58	HRLETSRSLASQIIGRPHSPEDVPLASVSPESMGPSA--ARNLRPNVGECPVF	115						
DB	61	HRLETSRPPASAEADIRRHSDYVAFASAT---GNPGLDPRAIRFVNGECPVF	116						
QY	116	DGLFPCRASAGSGISGAVALKNODADIAINMGGLHRAKKSASGFCVNDIVLGLLEL	175						
DB	117	DGLFPCQASAGSGISGAVALKNODADITVNNAGGCHHAKKSASGFCVNDIVLGLLEL	176						
QY	176	LKMFKRVLLIDIVHHGDEVEAFYTTDVMVYFHFKFDGFPFGTGHINDVCAEKAKYYA	235						
DB	177	LKHRRVLYVDIDVHHGDEVEAFYTTDVMVYFHFKFDGFPFGTGHINDVCAEKAKYYA	236						
QY	236	LNVPLNDGMDDESFRSLPPLLOKVMVEYOPAEVYVLOGGADSLSGORLCFNLISVGHAD	295						
DB	237	LNVPLSDGIDDTTFCGLFCQILKKVMEVYQPDVYVLOGGADSLSGORLCFNLISVGHAD	296						
QY	296	CLAFRLSYNVPMLVVLGEGEYTIIRNVARCMEYTAFAVGEVPPNKLIPYNYEFYFGPDYTL	355						
DB	297	CLAFRLSYNVPMLVVLGEGGTYIRNVARCMEYTAFAVGEVPPNKLIPYNYEFYFGPDYTL	356						


```
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,255A
FILING DATE: September 14, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terrence F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 482
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung cDNA library
US-08-528-255A-1

Query Match          56.9%; Score 1448.5; DB 1: Length 482;
Best Local Similarity 58.3%; Pred. No. 7.6e-152;
Matches 261; Conservative 79; Mismatches 81; Indels 27; Gaps 4;

QY 16 GRRKRVSYEPPTIGDYYGOGHPMKPHRIIMASLIHYHLRRLISRPSLADASDYG 75
DB 6 GTRKRVYVYDGVNYYGOGHPMKPHRIIMTHLLNTGLYRKMEIYRPHKANAEEMT 65

QY 76 RFRSPEYVDLASVSPESKMDPSAARLRRFNVEDCPYFDGLFDCRASAGSIGAAVK 135
DB 66 KYHSDYIKFLRSIRPDNMSEYS--KOMORFNVEDCPYFDGLFEFCQLSTGGSVASAVK 123

QY 136 LNRQDAIDAIINMGGLHHAKKSEASGFCYVNDIVLGLLELLKMKRYLYTIDIVHHGDV 195
DB 124 LNKQOTDIANNMAGGLHHAKKSEASGFCYVNDIVLLELLKMKRYLYTIDIVHHGDV 183

QY 196 EFAFYTTDRVMTVSFKFGDFPGTGHIRDVGAEKGYVALNVPYLDGMDESFSLFRP 255
DB 184 EFAFYTTDRVMTVSFKHYGEYFPGTGLRDIGAGKKYAVNPLRDGIDDESYEALFKP 243

QY 256 LIOKVMYVYDEAVYVLOCGASDLSGDRLGCPNLVSKGHACQRLRLRYNPLMWLGEGY 315
DB 244 VMSVYMEFQSAVYVLOCGSDLSGLDPCNLTIKHAACVEYVSKFNLPMMLLGGGGY 303

QY 316 TIRNVAQWCYETAVAAGVEPDNKLRYNEYFEGPYDYLHVDPSPENLTPKDMERIR 375
DB 304 TIRNVAQCRITYETAVALDTELPNLPYNDYFEGYRPFKHLISNSMNTQNTINYLEKIK 363

QY 376 NTLEQLSGLIHAPSVQFHTPPYNNRVLDEPDEDMETRPKPKXSGTAYTESDSDDDDKP 435
DB 364 QRLFENLRMLPHARGYVQMOAIPE-DAIPEESGDE-----DEDDDKR 404

QY 436 LHGYSRGATTDRTDSTGEDEMDNDPE 463
DB 405 I-----SICSSDKRIACEEFSDEE 426
```

```
RESULT 5
US-08-717-365-1
Sequence 1: Application US/08717365
Patent No. 5763182
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: FURUKAWA, YOICHI
TITLE OF INVENTION: RPOL PROTEIN AND DNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS, P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,365
FILING DATE: 23-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/528 255
FILING DATE: September 14, 1995
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terrence F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 482
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung cDNA library
US-08-717-365-1

Query Match          56.9%; Score 1448.5; DB 1: Length 482;
Best Local Similarity 58.3%; Pred. No. 7.6e-152;
Matches 261; Conservative 79; Mismatches 81; Indels 27; Gaps 4;

QY 16 GRRKRVSYEPPTIGDYYGOGHPMKPHRIIMASLIHYHLRRLISRPSLADASDYG 75
DB 6 GTRKRVYVYDGVNYYGOGHPMKPHRIIMTHLLNTGLYRKMEIYRPHKANAEEMT 65

QY 76 RFRSPEYVDLASVSPESKMDPSAARLRRFNVEDCPYFDGLFDCRASAGSIGAAVK 135
DB 66 KYHSDYIKFLRSIRPDNMSEYS--KOMORFNVEDCPYFDGLFEFCQLSTGGSVASAVK 123

QY 136 LNRQDAIDAIINMGGLHHAKKSEASGFCYVNDIVLGLLELLKMKRYLYTIDIVHHGDV 195
DB 124 LNKQOTDIANNMAGGLHHAKKSEASGFCYVNDIVLLELLKMKRYLYTIDIVHHGDV 183

QY 196 EFAFYTTDRVMTVSFKFGDFPGTGHIRDVGAEKGYVALNVPYLDGMDESFSLFRP 255
DB 184 EFAFYTTDRVMTVSFKHYGEYFPGTGLRDIGAGKKYAVNPLRDGIDDESYEALFKP 243
```


APPLICANT: MIYAKE, Kazuo
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Hiroyo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/446,504
CURRENT FILING DATE: 1999-12-23
PRIORITY APPLICATION NUMBER: PCT/JP98/02845
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: JP 9-187496
PRIORITY FILING DATE: 1997-06-26
PRIORITY APPLICATION NUMBER: JP 9-320692
PRIORITY FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 80
LENGTH: 335
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-446-504-80

Query Match 9.7%; Score 247.5; DB 4; Length 335;
Best Local Similarity 22.1%; Pred. No. 7.3e-19;
Matches 68; Conservative 64; Mismatches 139; Indels 37; Gaps 8;

QY 22 SYFEPTIGDYYGCGHPMKPHRIKMAHSLIIHYHLRLEISPSLADSDIGRFSPE 81
DB 4 SIFSKFKNLHRSPEYHDPENKRLIEIYSKVELGLEERIE--EPNVEEFVFKIKDRD 61
QY 82 YVDELASVSPESMGDPSAARNLRRFNVEDCPVFDGLFDFCRASAGSIGAAVNLNRDA 141
DB 62 YINVE-----AVEKIGITRLD--PDITYSPGTMSALLALGAARSAALS----- 104
QY 142 DIAINMG-----GLHAKKSEAS-----GFCYVNDIVIGILELKKM-FKRVLYID 186
DB 105 -ALHYGGIHALVRPGRHAGRRGRAMGAPTLGFCIFNNASAVVTLKEGVGKVVYID 162
QY 187 IDVHGGVEAEFTTDRVMYVSFHKFGDFRGTGHIRDVAEKGKYYALNPLNDGMD 246
DB 163 FDAHNGMGTQIFNDDPDVYIHIDHE--RDYIPGSGDVSEGGSNAYSKIMLPMHYSGD 221
QY 247 ESFSLFPLIQLQKMEVYQPAVVLQCGADSLSGDRLCFNLVSKGHADCLRFKSYNP 306
DB 222 GDYIYVNDIYLPVVEEKRVYIYISAGFDGFKDGLTLRLTLETSFYAGATLTKIPIA 281
QY 307 LMLVGGES 314
DB 282 FILEGGYG 289

RESULT 9
US-09-144-085-2
Sequence 2, Application US/09144085
Patent No. 6280999
GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziemann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 30062-20020.20
CURRENT APPLICATION NUMBER: US/09/144,085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010,809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 6095

TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match 4.1%; Score 104.5; DB 4; Length 6095;
Best Local Similarity 20.1%; Pred. No. 0.68;
Matches 109; Conservative 59; Mismatches 193; Indels 181; Gaps 24;

QY 2 EADESGISLSPGDPGKRRVSYFEPTIGDYGGGHPMKPHRIKMAHSLIIHYHLRRI 61
DB 1426 ELDPAGTGLVLTGTGE-----LQALIAHLVR-AHG-VRLVLTLSRR 1465
QY 62 FISRPSLADSDIGRFSPEYVDFLASVSPESMG-----DPSAARNLRRFNVED 111
DB 1466 GLEAPG-----ARLVOSLEELGAETVVAACDVSKREEARVLAGIDAPLPS 1514
QY 112 -----CPVPDG--LFDGR-----ASAGSGTGA 132
DB 1515 AVHLIAGVLDGVLTAOTARLSRYLAPKVDGALHLHLRELDLAAFLVSSAGTPGA 1574
QY 133 AVKLNR-----DADIAINMGGLHAKKSEASGFCYVNDI---VLGILELKKFKRY 182
DB 1575 AGOSNYAANTFLDALAAHRRGGGL--AATSLAMGEFTQAGVOMTAHIGFAELSRM-RRM 1631
QY 183 LYIDIDVHGDVGEAETTDRTVMYVSFHKFGDFRGTGHIRDVAEKGKYYALNPLND 242
DB 1632 GFVPMPEESGLALIDALASPEASLYPVR-----IDL-----AQLR 1668
QY 243 GMDDES-----FSLFPLIQLQKMEVYQPAVVLQCGADSL--SGDRLCFNLVSKGHAC 296
DB 1669 GLESSGELPALFALLRLPSLRKSSAATRRDASLRELIALPAERLNLVLYREYVA 1728
QY 297 LRFRLSYNPVLMVLGGGCTITRVARQWCYETIYVANGV-----EPDNKLPVNEYEYEG 350
DB 1729 VAGLQR-----GEAVALDVLYKELGIDSLAAVALRNRLTSRTETSLPATVFDY 1777
QY 351 PDYTLVHDPSP-----MENLTPKDMERIRN-----TLLEGLSLIHA 388
DB 1778 -----PFRALIAELLQAFSGLOVKERARARRACKDEPIAIVASACLPGGVAT 1829
QY 389 PSVOFQHTPRVNRVLDPEPDMETRPKRXWSGTATYESDDDDKPLHGYSCRGATTD 448
DB 1830 PDDVW-----RLLAGKDAIE--GIIPARMDGFEVYDPPPEAAGK---SYAREGGFVRD 1877
QY 449 RD 450
DB 1878 ID 1879

RESULT 10
US-09-428-517-2
Sequence 2, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence

FILED DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-867-611-2

Query Match 3.6%; Score 91; DB 4; Length 396;
Best Local Similarity 22.4%; Pred. No. 0.22; Mismatches 83; Indels 88; Gaps 14;

Matches 59; Conservative 33; Mismatches 83; Indels 88; Gaps 14;
194 GVEEAFYTR-----VMTVSFKGDFPGTGHIRDVGAEKGYVALNP 239
41 GAERIVATDIEDVARAVEAAGEVCHTRADQ-----STERIAEV-VEKCAFSDDTVI 94
240 LNDGMDSEFSRLFRPLIOKMEVYQPEAVYVLOCADSLSGDLGCFNLVSKGHADCLRF 299
95 VNVQGD-----PMI-----PATIRQV-ADNLAQGVGMATLAVPIHNAEAF 137
300 LRSNVPLMLVGGEGTIRNVARCMCYETAAVGVPEPDNKLPINEYFPGPYTHV-- 357
138 --NPNNAKVVLDAGYAL-----YFSRATIPWDRD--FAEGLETYGDNFLRLHGI 184
358 -----DPSMENLTPKDMERIRNTLLEQLSGLIHAPSVQF---QHTPPVN 400
185 YGYRAGTIRRYVNMQPSPLEHIE-----MLEQLRVLMYGEKIHVAVAQVEPGTG 233
401 RVLDEPED-----DMETRPKPR 417
234 --VDTPELDLPSTNSMSTNPKPQ 254

RESULT 13
PCT-US92-06965A-7
Sequence 7, Application PC/TUS9206965A
GENERAL INFORMATION:
APPLICANT: DEVAR, S.
APPLICANT: DAILEY, S.
TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: ILLINOIS
COUNTRY: U.S.
ZIP: 60065-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06965A
FILING DATE: 19920821
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834PC.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-06965A-7

Query Match 3.6%; Score 91; DB 5; Length 396;
Best Local Similarity 22.4%; Pred. No. 0.22; Mismatches 83; Indels 88; Gaps 14;
Matches 59; Conservative 33; Mismatches 83; Indels 88; Gaps 14;

194 GVEEAFYTR-----VMTVSFKGDFPGTGHIRDVGAEKGYVALNP 239
41 GAERIVATDIEDVARAVEAAGEVCHTRADQ-----STERIAEV-VEKCAFSDDTVI 94
240 LNDGMDSEFSRLFRPLIOKMEVYQPEAVYVLOCADSLSGDLGCFNLVSKGHADCLRF 299
95 VNVQGD-----PMI-----PATIRQV-ADNLAQGVGMATLAVPIHNAEAF 137
300 LRSNVPLMLVGGEGTIRNVARCMCYETAAVGVPEPDNKLPINEYFPGPYTHV-- 357
138 --NPNNAKVVLDAGYAL-----YFSRATIPWDRD--FAEGLETYGDNFLRLHGI 184
358 -----DPSMENLTPKDMERIRNTLLEQLSGLIHAPSVQF---QHTPPVN 400
185 YGYRAGTIRRYVNMQPSPLEHIE-----MLEQLRVLMYGEKIHVAVAQVEPGTG 233
401 RVLDEPED-----DMETRPKPR 417
234 --VDTPELDLPSTNSMSTNPKPQ 254

RESULT 14
US-08-291-896-2
Sequence 2, Application US/08291896
Patent No. 5480800
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA fragment carrying the gene encoding the
TITLE OF INVENTION: enzyme for fragmenting N-acetyltheparosan and the adjacent
sequences permitting its expression, recombinant enzyme
TITLE OF INVENTION: intended for fragmenting N-acetyltheparosan and its use.
NUMBER OF SEQUENCES: 7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (Epo)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,896
FILING DATE: 17-AUG-1994
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 820 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-291-896-2

Query Match 3.58; Score 88; DB 1; Length 820;
Best Local Similarity 19.88; Pred. No. 1.6;
Matches 69; Conservative 36; Mismatches 123; Indels 120; Gaps 17;

QY 164 YVNDIVLGILELLMKFRVLYIDIDVHNGGVEAEATYTDKRVMTVSFHNFGDFPFGCHI 223
DB 163 YVDAALAGVIGFGWF-----IYSGS---AVYRT-----FQDKMRDGVSI 200
QY 224 RDVGAEGKYYALNVLPLDGMDESFSLRPLIOKMEVYOEAV-----VLQCGAD 276
DB 201 KDFGAQNG-----LINDNKD-----AFTKSLHSFSSVFVEGFYNTSLVSLRGLY 247
QY 277 SLSG-----DRLG---CFNL-----SYKHADCLRLRSYVPLMV 309
DB 248 GTGGGTIKQYDRDGNHLVFNMPDGMSTLTIMGKNSDSDSVQGH-----QVSF 295
QY 310 LGSGGTIRNVARCMCEYTAAYGVEPDNKL-----YNEVFY-----FGP 351
DB 296 SGGHDVSKNIRFTNTRGPGFSLIAYPDNGIPSGYIVRDIRGEYLGFAANNKACVLPDS 355
QY 352 DYTLLHVPSPMEN-----LNTPKDMERIRNTLLEQLSLIHAPSVOFQHTPPVNRVL 403
DB 356 SQNTLLDGVIAIRNYPQGAVELKTAAKYNIIVSNVIGECQHVYNGTE--TETAPNTNII 413
QY 404 DEPEDMETRPKPRX---WSGTATYESD-----SDDDKPLHGYSCRG 443
DB 414 SSV---MANNPKYAAVYVKGCTGNLISDVLVDYSESDAKAHGVTVOG 458

RESULT 15

US-08-485-278-2

Sequence 2, Application US/08485278
Patent No. 5820857
GENERAL INFORMATION:
APPLICANT: Legoux, Richard
APPLICANT: Lelong, Philippe
APPLICANT: Salome, Marc Louis Victor
TITLE OF INVENTION: DNA fragment carrying the gene encoding the enzyme for
TITLE OF INVENTION: fragmenting N-acetylheparosan and the adjacent sequences perm
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern
STREET: 400 Seventh St. N.W.
CITY: Washington D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,278
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93/10050
FILING DATE: 17-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/291,896
FILING DATE: 17-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: P58114NA

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-5350
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 820 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-278-2

Query Match 3.58; Score 88; DB 2; Length 820;
Best Local Similarity 19.88; Pred. No. 1.6;
Matches 69; Conservative 36; Mismatches 123; Indels 120; Gaps 17;

QY 164 YVNDIVLGILELLMKFRVLYIDIDVHNGGVEAEATYTDKRVMTVSFHNFGDFPFGCHI 223
DB 163 YVDAALAGVIGFGWF-----IYSGS---AVYRT-----FQDKMRDGVSI 200
QY 224 RDVGAEGKYYALNVLPLDGMDESFSLRPLIOKMEVYOEAV-----VLQCGAD 276
DB 201 KDFGAQNG-----LINDNKD-----AFTKSLHSFSSVFVEGFYNTSLVSLRGLY 247
QY 277 SLSG-----DRLG---CFNL-----SYKHADCLRLRSYVPLMV 309
DB 248 GTGGGTIKQYDRDGNHLVFNMPDGMSTLTIMGKNSDSDSVQGH-----QVSF 295
QY 310 LGSGGTIRNVARCMCEYTAAYGVEPDNKL-----YNEVFY-----FGP 351
DB 296 SGGHDVSKNIRFTNTRGPGFSLIAYPDNGIPSGYIVRDIRGEYLGFAANNKACVLPDS 355
QY 352 DYTLLHVPSPMEN-----LNTPKDMERIRNTLLEQLSLIHAPSVOFQHTPPVNRVL 403
DB 356 SQNTLLDGVIAIRNYPQGAVELKTAAKYNIIVSNVIGECQHVYNGTE--TETAPNTNII 413
QY 404 DEPEDMETRPKPRX---WSGTATYESD-----SDDDKPLHGYSCRG 443
DB 414 SSV---MANNPKYAAVYVKGCTGNLISDVLVDYSESDAKAHGVTVOG 458

Search completed: April 26, 2002, 17:51:27
Job time: 35 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:13 ; Search time 43 Seconds
(without alignments)
811.361 Million cell updates/sec

Title: US-09-645-337-4

Perfect score: 2544
Sequence: 1 MEADSSGSLSPSGPDCRRKRR.....TGDEMDDDNPEPDVNPSS 471

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A-Geneseq-1101:*

1:	/SID58/gcgdata/geneseq/geneseq/AA1980.DAT:*
2:	/SID58/gcgdata/geneseq/geneseq/AA1981.DAT:*
3:	/SID58/gcgdata/geneseq/geneseq/AA1982.DAT:*
4:	/SID58/gcgdata/geneseq/geneseq/AA1983.DAT:*
5:	/SID58/gcgdata/geneseq/geneseq/AA1984.DAT:*
6:	/SID58/gcgdata/geneseq/geneseq/AA1985.DAT:*
7:	/SID58/gcgdata/geneseq/geneseq/AA1986.DAT:*
8:	/SID58/gcgdata/geneseq/geneseq/AA1987.DAT:*
9:	/SID58/gcgdata/geneseq/geneseq/AA1988.DAT:*
10:	/SID58/gcgdata/geneseq/geneseq/AA1989.DAT:*
11:	/SID58/gcgdata/geneseq/geneseq/AA1990.DAT:*
12:	/SID58/gcgdata/geneseq/geneseq/AA1991.DAT:*
13:	/SID58/gcgdata/geneseq/geneseq/AA1992.DAT:*
14:	/SID58/gcgdata/geneseq/geneseq/AA1993.DAT:*
15:	/SID58/gcgdata/geneseq/geneseq/AA1994.DAT:*
16:	/SID58/gcgdata/geneseq/geneseq/AA1995.DAT:*
17:	/SID58/gcgdata/geneseq/geneseq/AA1996.DAT:*
18:	/SID58/gcgdata/geneseq/geneseq/AA1997.DAT:*
19:	/SID58/gcgdata/geneseq/geneseq/AA1998.DAT:*
20:	/SID58/gcgdata/geneseq/geneseq/AA1999.DAT:*
21:	/SID58/gcgdata/geneseq/geneseq/AA2000.DAT:*
22:	/SID58/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2542	99.9	471	22	AA67812	Amino acid sequenc
2	2535	99.6	471	21	AA605187	Arabidopsis thalia
3	2535	99.6	471	21	AA639372	Arabidopsis thalia
4	2318	91.1	432	21	AA605188	Arabidopsis thalia
5	2318	91.1	432	21	AA639373	Arabidopsis thalia
6	2279	89.6	425	21	AA605189	Arabidopsis thalia
7	2279	89.6	425	21	AA639374	Arabidopsis thalia
8	1734	68.2	458	20	AA128797	Maize histone deac
9	1561.5	61.4	501	22	AA67811	Amino acid sequenc
10	1561.5	61.4	501	22	AA67165	Arabidopsis thali
11	1558.5	61.3	473	21	AA58829	Soybean histone de

12	1551.5	61.0	517	20	AA28800	Maize histone deac
13	1520	59.7	439	20	AA28799	Maize histone deac
14	1498	58.9	493	21	AA58828	Rice histone deace
15	1462.5	57.5	482	18	AAW29324	A novel histone de
16	1462.5	57.5	482	22	AA67164	Human histone deac
17	1462.5	57.5	482	22	AA64954	Human histone deac
18	1462.5	57.5	488	21	AA65685	Human histone deac
19	1448.5	56.9	482	17	AA68819	Human prostate can
20	1445.5	56.8	488	22	AA64955	Human foet
21	1348.5	53.0	430	22	AAU00243	Human histone deac
22	1328.5	52.2	433	22	AA67166	Human histone deac
23	1292.5	50.8	428	22	AA64956	Human histone deac
24	1280.5	50.3	351	20	AA28798	Human histone deac
25	1249.5	49.1	432	20	AA28801	Maize histone deac
26	813.5	32.0	415	22	AAW40374	Human histone deac
27	813	32.0	377	21	AAW83027	Human histone deac
28	813	32.0	377	22	AA64361	Human histone deac
29	594	23.3	180	22	AA670496	Amino acid sequenc
30	561.5	22.1	296	22	AAW42160	Saccharomyces cere
31	465.5	18.3	375	22	AAU00241	Human polypeptide
32	460	18.1	125	18	AAW29330	Cys75Ser/Cys77Ser
33	459	18.0	133	18	AAW29329	A partial deacetyl
34	457.5	18.0	375	22	AAU00240	Wild-type histone
35	453.5	17.8	375	22	AAU00242	Tyr297Phe mutant o
36	328.5	12.9	120	21	AAW58830	Wheat histone deac
37	309.5	12.2	74	22	AA648580	Human ovarian can
38	309	12.1	69	18	AAW29326	Hox protein family
39	309	12.1	398	21	AA607246	Arabidopsis thalia
40	302	11.9	364	21	AA607248	Arabidopsis thalia
41	302	11.9	377	21	AA607247	Arabidopsis thalia
42	295	11.6	80	18	AAW29331	A partial deacetyl
43	289	11.4	468	21	AA651592	Arabidopsis thalia
44	289	11.4	499	21	AA651591	Arabidopsis thalia
45	289	11.4	552	21	AA651590	Arabidopsis thalia

ALIGNMENTS

RESULT	1
ID	AA67812
AA67812	standard: Protein; 471 AA.
AC	AA67812:
DT	29-JUN-2001 (first entry)
XX	
DE	Amino acid sequence of a histone deacetylase designated AtRPD3B.
XX	
KW	Histone deacetylase: AtRPD3B; RPD3; gene expression: transgenic plant;
KW	HDA1; ethylene-responsive phenotype: hypocotyl elongation.
XX	
OS	Arabidopsis thaliana.
XX	
PN	CA2316036-A1.
XX	
PD	27-FEB-2001.
XX	
PF	24-AUG-2000; 2000CA-2316036.
XX	
PR	27-AUG-1999; 99US-0383971.
XX	
PA	(MTC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX	
PI	MIKI B, Brown D, Tian L, Wu K;
XX	
DR	WPI: 2001-258457/27.
XX	
DR	N-PSDB; AAF60351.
XX	
PT	Methods for regulating gene expression in transgenic plants, e.g.
PT	repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT	elongation), comprises introducing genes encoding histone deacetylase

XX Claim 10; Fig 1B; 91pp: English.
PS
XX
CC The present sequence represents Arabidopsis thaliana histone deacetylase
CC designated AtRpd3B. The protein is homologous to yeast Rpd3 and HDAC1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
SO Sequence 471 AA:

Query Match 99.9%; Score 2542; DB 22; Length 471;
Best Local Similarity 99.8%; Pred. No. 6.2e-263;
Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEADESGISLPSPGDKRRVSYTEPTIGDYGGCHPMKPHIRMAHSLIHYHLHRR 60
Db 1 meadesgslspspgdkrrvsvteptigdyggchpmkphirmahsllhyhlhrr 60

QY 61 LEISRPLADSDIGRFRSPRYVDPLASVSPESMGDPSAARNLRFWGDCPPEDGLFD 120
Db 61 leisrpladadslgrfrhspeyvdlasvspesmgdpsaarnlrirvngedcpfdglfd 120

QY 121 PCRASAGSGICGAAYKLNKODADIAINNGGGLHNAKSEASGFCYVNDIVGILELLKMF 180
Db 121 pcraasagsgicgaayklnkoadadialnngggllhnaakseasgfcyvndivgillellkmfk 180

QY 181 RVLYIDIDVHNGDVEAFYTTDVMKTVSFHKKDFPFGTGHIDVCAEKGKITYALVPL 240
Db 181 rvlyididvhnngdveaftyttvmkvtvsfhkkdfpfgtghidvcaeakgkityalvpl 240

QY 241 NDGMDDESRFLRPLLOKVMVEYQPEAVVYLCQCADSLSGRLCFNLNVKAGHDCRLFL 300
Db 241 ndgmddeesrflrpllokvmveyqpeavvylcqcadslsgrlcfnlsvkaghdcrlfl 300

QY 301 RSYVPLMLVIGEGYTRRNARCMCYETAAVAVGEPDNKLRYNEFEYFGPDYTLHVDPS 360
Db 301 rsyvplmlvlgegytrrnarcmcycetavaavgepdnklrynefeyfgpdytlhvdp 360

QY 361 PMENLNPCKMERIRNLLLEGLIHAPSVQFQRTFPRVNRVLDPRDDMETRKPXMS 420
Db 361 pmenlnpkmerirnlleegllhapsvqfqrtpfvnrvidprddmetrpkpms 420

QY 421 GTATYESDDDDDKPLAGYSGRCGATTDSDTGEDMDODDPEPDVNPSS 471
Db 421 gtatyesdddddplagysgrcgattddstggedmdoddpepdvnpss 471

RESULT 2
ID AAG05187 standard; Protein; 471 AA.
XX AAG05187;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1499.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.

XX
PN EPI033405-A2.
XX
PU 06-SEP-2000.
XX
PR 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 28-APR-1999; 99US-0130891.
XX
PR 30-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 04-MAY-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 14-MAY-1999; 99US-0134256.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 18-MAY-1999; 99US-0134370.
XX
PR 19-MAY-1999; 99US-0134461.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135153.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136392.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.
XX
PR 14-JUN-1999; 99US-0139119.
XX
PR 16-JUN-1999; 99US-0139452.
XX
PR 16-JUN-1999; 99US-0139453.
XX
PR 17-JUN-1999; 99US-0139452.
XX
PR 18-JUN-1999; 99US-0139454.
XX
PR 18-JUN-1999; 99US-0139455.
XX
PR 18-JUN-1999; 99US-0139456.
XX
PR 18-JUN-1999; 99US-0139457.
XX
PR 18-JUN-1999; 99US-0139458.
XX
PR 18-JUN-1999; 99US-0139459.
XX
PR 18-JUN-1999; 99US-0139460.
XX
PR 18-JUN-1999; 99US-0139461.
XX
PR 18-JUN-1999; 99US-0139462.
XX
PR 18-JUN-1999; 99US-0139463.
XX
PR 18-JUN-1999; 99US-0139750.
XX
PR 18-JUN-1999; 99US-0139753.
XX
PR 21-JUN-1999; 99US-0139817.
XX
PR 22-JUN-1999; 99US-0139899.
XX
PR 23-JUN-1999; 99US-0140353.
XX
PR 23-JUN-1999; 99US-0140354.
XX
PR 24-JUN-1999; 99US-0140695.
XX
PR 28-JUN-1999; 99US-0140823.
XX
PR 29-JUN-1999; 99US-0140921.
XX
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 21-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 26-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.

PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Query Match 99.6% Score 2535; DB 21: Length 471.
Best Local Similarity 99.6% Pred. No.3.5e-262;
Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEADSGISLPSPGDDKRRVSEYFPTIGDYVYCGGHPMKPHRIIMASLIITHYLR 60
|||||
Db 1 meadsgislpSPGDDKRRVSEYFPTIGDYVYCGGHPMKPHRIIMASLIITHYLR 60
|||||
QY 61 LEISRPSLADSDTGRFHSBEYVDPLASVSPESMGDPSAARNLRPNVGEDCPVDFGLFD 120
|||||
Db 61 leisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrpnvgedcpvdfglfd 120
|||||
QY 121 FCRASAGSISGAAVKINRDOADIAIMWGGGLHRAKKEASGFCYVNDIYLGILLEIKMKR 180
|||||
Db 121 fcrasagsisgaavkinrqadaiaimwggglhakrakseasgfcyvnidiylgilleikmk 180
|||||
QY 181 RVLVYIDIDVHHGDDVEAEFTTDRVMTVSGFHKRGDFPCTGHTRDVGAERGYALNVPL 240
|||||
Db 181 rvlvlyididvhgddveaefttdrvmtvsgfhtkrgdfpctghtrdvgaekgyalnvpl 240
|||||
QY 241 NQGMDEFSRSLFRPLIQVMEYVQPEAVVLCGADSLSGDRIGCENISVKHADCLRFL 300
|||||
Db 241 nqgmdefsrslfrpliqvmevyqpeavvlcgadslsgdrigcfnlsvkhadclrfl 300
|||||
QY 301 RSYNVPLMWLGCGETIRNVARCMQETAAVAGVPEPDNKLRYNEVEYEGPYTLTHVDS 360
|||||
Db 301 rsynvplmwlgcgetirnvarcmqetAAVAGVPEPDNKLRYNEVEYEGPYTLTHVDS 360
|||||
QY 361 PMENLTPKDMERIRNTLLEOLSGLIHAPSVOFQHPNPPVNRVLDEPDMETPRPKRMS 420
|||||
Db 361 pmenltpkdmerrirntllleolsglihapsvofqhpnpvnrvldepdmetprprkms 420
|||||


```

PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 25-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151430.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 28-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158332.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161921.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 99.68; Score 2535; DB 21; Length 471;
 Best Local Similarity 99.68; Pred. No. 3.5e-262;
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MEADESCISLPSGPGKRRVSYEPTIGDYVGCGHPKPHIRRAHSLIIHYHLRR 60
Db 1 meadeslslpsgpgdkrrvsyfeptldgyygghpmkphirrahsllihyhllrr 60
OY 61 LEISRPISLADSDIGRPHSPFYVDLFASVSPESMGDPSAARNLRRFNVGDCPVFDGLFD 120

```

```

Db 61 leisrpladadigrfhspeydfiasvpsesmgdpsaarnlrrfnvgdcpvfdglfd 120
OY 121 FCRAAGSGTGAAYKLNRODADIAIMWGGGLHAKKSEASGCEYVNDIYLGILLELKKFX 180
Db 121 fcraasgsgtgaayklnroqadialmwggglhaksseasgicyndivlgillelkkfx 180
OY 181 RVLVYIDVHAGDVEAEFTTDRNMTVSFHKRGDFPCTGHIROYGAKGKYALNVL 240
Db 181 rvlvldvvhagdvaeafyldrvmtvsfhrkgdfpctgthirvgaekgyyalnvl 240
OY 241 NDGMDESPRSIFRPLIOKMEVYOPPAVLIQCADSLSGDRICFNLSVKHACDLRFL 300
Db 241 ndgmdestrsifrlplioqkmevygpaavliqgadslsgdrigcfnlsvkhacdlrfl 300
OY 301 RSYNPLMLVLCGEGYTIIRNARWCYETAVAVGVEPDNKLPIREYREYGPOTYTLHVDPS 360
Db 301 rsynplmlvlggytirnvarwcycetavavvepdnklpiyreyletgyptylhvdps 360
OY 361 PMENLTPKDMERIRNTLLEQLSLIHAPSVOFQHTPPNRYLDEPEDDMETRPKPRXS 420
Db 361 pmenltpkdmertirntllleqslilhapsvqfhtppnryldepeddmctrpkprxs 420
OY 421 GTATYESDSDDDDKPLHGYSCNCGATTDRDSTGEDMDNDPEPDVNPSS 471
Db 421 gtatyesdsdddkplhgyscrgcattdrdsigedemddndpepdvnpss 471

RESULT 4
ID AAG05188 standard; Protein: 432 AA.
XX
AC AAG05188;
XX
DE 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1500.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-01231825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.

```

PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 14-JUN-1999; 9905-0138847.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140891.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.

PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152463.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 91.1%; Score 2318; DB 21; Length 432;
Best Local Similarity 99.5%; Pred. No. 5; Le-239;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 40 MKPHRIRIMASLIIRHYLHRLRLSRLSLADSDIGRHSPEYVDLIVLASVSPESKDPVSA 99
DB 1 mkphritmahslilhyhlhrrlrlslsladadigrhspeyvdlivlasvpsesmgdpsaa 60
QY 100 ARNLRRFNVGECDEPVEFGDLDFPCRASAGSICAAVKNLRDADIAINMGCGLHNAKSPA 159
DB 61 arnlrrfnvgedcpvldgldfdrctasagsgaavknrddadialnwgqghlhaaksee 120
QY 160 SGFCYVNDIVLGILIELIKMKRVLVIDIDVHNGDGYEAEYTTDRVWVVSFHKGGDFPG 219
DB 121 sgfcyvndivlgiliekmlkrrvlyididvhdggyeaeayttldrmvtsfhkfgdfpg 180
QY 220 TGHTRGVGAERKGYVALNVLNDCMDDESFRSLFRLIOLKVMVNVQPREANVLQCGADSL 279
DB 181 tghtrdvgaekgyvalnvplndgmddesfrslfrplqlkvmevyppeavvlqcgadsis 240
QY 280 GDRIGCFNLISVKGHADCLIRLRSYNPLWVLAGEGYTIRRVARCKWCEYFAVAAGVFPDK 339
DB 241 gdrigcfnlisvkghadclirfrrsynpwlwlggygylrrnarckwceyetaavagvfpdk 300
QY 340 LPYNEYFEYFGPDYTLAHVDSPEMENLTPKDMERIRNTLLEQLSLIHAPSVQFHTPPV 399
DB 301 lpyneyfeyfgpdytlhvdpsmenltpkdmerrntllleqslghapsvqfhtlppv 360
QY 400 NRVLDPEDDMETPRKRXMSGATYESDSDDDKPLHGSCHSGAGATTRODSRGDEMD 459
DB 361 nrvldepddmetprkprlwsigatyesdoddckprlhyscrgrgatldrdsrgedemd 420
QY 460 DNPEPDVNPSS 471
DB 421 dnpepdvnpss 432

RESULT 6
AACG05189
ID AACG05189 standard; Protein; 425 AA.
XX

AC AAG05189;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1501.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125768.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134286.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139495.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148865.
PR 13-AUG-1999; 99US-0148864.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139452.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 21-JUN-1999; 9905-0139763.
PR 22-JUN-1999; 9905-0139817.
PR 23-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142055.
PR 08-JUL-1999; 9905-0142390.
PR 09-JUL-1999; 9905-0142803.
PR 12-JUL-1999; 9905-0142920.
PR 13-JUL-1999; 9905-0142977.
PR 14-JUL-1999; 9905-0143542.
PR 15-JUL-1999; 9905-0143624.
PR 16-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144612.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.

PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148365.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151308.
PR 01-SEP-1999; 9905-0151330.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 24-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157153.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160747.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 89.6%; Score 2279; DB 21; Length 425;
 Best Local Similarity 99.5%; Pred. No. 7,4e-235;
 Matches 423; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 47 MAHSLIIHHHRLLEISRSLSLADSDIGRHSPEYVDPLASVSPESMGDPSAARNLR 106
 DB 1 mahslilhyhrrleisrslasdasdigrhspeyvdilassvpsmgdpsaarnlr 60
 QY 107 NVGDCVPFDGLFDFCRASAGSIGAAYVKLRQADAIINMGGLHHAKKSEASGFCYVN 166
 DB 61 nvgedcpvifdglfdcrasagsgigaavklnrqadaiainmggslhakkseasgfcyvn 120
 QY 167 DIVIGILELLKMKRVLVYIDIVHHGDCVEAFYTTDRVMTVSFHKFGDFPGTGHIRDV 226
 DB 121 divigilellkmkfrvlyidivhghgveaftytdrvmtvsfhkfgdfpyghirdv 180
 QY 227 GAEGKAYVALNVRPLNDGMDSEFSRLFRPLIQKMEYVOPFAVVLQCGASLSGDRIGCF 266
 DB 181 gaekayvalnvrplndgmdsefsrlfrpliqkmevyopfaavvlqcgaslsdgdrigcf 240
 QY 287 NLAYKGNADCLRLRSYNNVPLMVLGEGYTIIRNVARCMCYETAAVAVGEPPDNKLPYNEYF 346
 DB 241 nlavkgnadclrlrsynnvplmvlgggytliirnavarcwcyetavaavgeppdnklpyneyf 300
 QY 347 EYFGPDYTLHAVDPSPMENLTPKDMERIRNTLLEQLSGLIHAPSVOFQHTPPVNRVLDEP 406
 DB 301 eyfgpdytlhavdpspmenltpkdmerrntlleqlsghlhapsvofqhtppvnrvldep 360
 QY 407 EDDMETPRKRWGSGTATYESDSDDDDKPLHGYSCRGGATTDROSTGDEMDNDNPEPDV 466
 DB 361 eddmctrprkrlwsgtatesdsdddkplhgyscrggattdrostgedemndnpepdv 420
 QY 467 NPSS 471
 DB 421 npss 425

RESULT 8
 ID AAY28797 standard; protein: 458 AA.
 XX
 AC AAY28797;
 XX
 DT 13-JAN-2000 (first entry)
 XX
 DE Maize histone deacetylase-1.
 XX
 KM Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
 KM acetyl modification; promoter; regulatory element; transgenic plant;
 KM disease resistance; toxin screening; pathogenicity;
 KM disease response promoter.
 XX
 OS zea mays.
 XX
 PN W0951731-A2.
 XX
 PD 14-OCT-1999.
 XX
 PE 02-APR-1999; 99MO-US07370.
 XX
 PR 03-APR-1998; 98US-0080563.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Baldwin DA, Briggs SP, Crane VC;
 XX
 DR WP1: 1999-611038/52.
 XX
 DR N-PSDB: AAX90837.
 XX
 PT New deacetylase genes, used for producing transgenic plants which have
 PT increased disease resistance -

PS Claim 1; Page 49-51; 87pp; English.

XX The present sequence is maize histone deacetylase encoded by HD cDNA
 CC belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
 CC modifications, may be localised to promoters targeted for repression by
 CC other proteins that associate with HD and specifically bind regulatory
 CC elements in promoter DNA. The HD nucleotide sequence can be used for
 CC producing transgenic plants with increased disease resistance.
 CC Additionally, compositions find use in screening for toxins that affect
 CC pathogenicity and in determining which disease response promoters are
 CC regulated by histone deacetylase.

SO Sequence 458 AA;

Query Match 68.2%; Score 1734; DB 20; Length 458;
 Best Local Similarity 69.3%; Pred. No. 1.8e-176;
 Matches 328; Conservative 50; Mismatches 73; Indels 22; Gaps 7;

QY 1 MEADESGISLPS--GPDGRKRVSYFEPTIGDYVYGQCHPMKPRIRMAHSLIIHHH 57
 DB 1 maesgslpspaggedahrrvsvfyfepsigdyyyggghpmkprirmahslvhy91 60
 QY 58 HRLLEISRSLSLADSDIGRHSPEYVDPLASVSPESMGDPSA--ARNLAKPWNCECPVF 115
 DB 61 hrleisrslasdasdigrhspeyvdplassvpsmgdpsa--arnlarkpwngeccpvf 116
 QY 116 DGLFDCRAGAGSISGAAYVKLRQADAIINMGGLHHAKKSEASGFCYVNDIVLGILEL 175
 DB 117 dglfrcgasaagsigaavklnrqadaitnmggslhakkseasgfcyvndivlailel 176
 QY 176 LKMKRVLVYIDIVHHGDCVEAFYTTDRVMTVSFHKFGDFPGTGHIRDVGAEGKAYVA 235
 DB 177 lkmkfrvlyidivhghgveaftytdrvmtvsfhkfgdfpyghirdvgaekayva 236
 QY 236 LNPVLNDGMDSEFSRLFRPLIQKMEYVOPFAVVLQCGASLSGDRIGCFNLSTVYGHAD 295
 DB 237 lnpvlsgidldtfrglfrcilkkmevyopfaavvlqcgaslsdgdrigcfnlstvyghad 296
 QY 296 CLRLRSYNNVPLMVLGEGYTIIRNVARCMCYETAAVAVGEPPDNKLPYNEYFXYFGPDYTL 355
 DB 297 clrlrsynnvplmvlgggytliirnavarcwcyetavaavgeppdnklpyndyefyfgpdytl 356
 QY 356 HVDPSPMENLTPKDMERIRNTLLEQLSGLIHAPSVOFQHTPPVNRVLDEPEDMETRP- 414
 DB 357 hvpksvenltpkdmerrntlleqlsghlhapsvofqhtppvnrvldepdmektrp 416
 QY 415 -KPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDROSTGDEMDNDNPEPDV 466
 DB 417 qsrflwsga-ydsdtedpsl-----ksqkdvlanlqmke-pkddl 458

RESULT 9
 ID AAB67811 standard; protein: 501 AA.
 XX
 AC AAB67811;
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE Amino acid sequence of a histone deacetylase designated AtRPD3A.
 XX
 KM Histone deacetylase; AtRPD3A; RPD3; gene expression; transgenic plant;
 KM HD01; ethylene-responsive phenotype; hypocotyl elongation.
 XX
 OS Arabidopsis thaliana.
 XX
 PN CA2316036-A1.
 XX
 PD 27-FEB-2001.
 XX
 PF 24-AUG-2000; 2000CA-2316036.
 XX


```

QY 304 NPLMLVGGEGYTIIRNVAKWCYETAVAVGVEDPNKLPYNEFEYFGPDYTLHVDSPME 363
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 nvp11119999ytlrnavarcwctetsvalgieddkmpqheyeyfgpdytlhvapsme 359
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 364 NLNTPKDMERIRMTLLEQLSLIHAPSVOGHTPPVNRV--LDEPEDMETRPRKPRXMSG 421
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 nksrq11delrak11dnslrlqhapsvpfqerppdterpevedgdkrwddps--- 416
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 422 TATYESDSDDDKPR-----LHGYSCHG-CATTPDRDSTGEDMDNDN 461
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 417 ----dmvdddrkp1psrvkreavepdkdkg1kqjmergkqcevedsgstkvtgyn 472
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 462 P 462
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 473 P 473
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
AAV58829 standard; Protein: 473 AA.
ID AAV58829
AC AAV58829;
XX 08-MAY-2000 (first entry)
DT
DE Soybean histone deacetylase 1 (HD1) protein.
XX
KM Chromatin associated protein; histone deacetylase gene 1; HD1:
KW soybean; transgenic plant; transcription regulation.
XX
OS Glycine max.
XX
MO200004177-A1.
XX
PN 27-JAN-2000.
PD
PF 13-JUL-1999; 99MO-US15807.
XX
PR 14-JUL-1998; 98US-0092841.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Vollmer SJ;
XX
DR MPI; 2000-182439/16.
DR N-PSDB; AAZ58260.
XX
XX New nucleic acid fragment useful as probes and primers, for
PT transforming plants
PS
PS Claim 1; Page 27-28; 36pp; English.
XX
XX The present sequence is that of soybean histone deacetylase 1
CC (HD1), a chromatin associated protein, as deduced from a soybean
CC root cDNA clone (see AAZ58260) isolated on the basis of homology to
CC plant histone deacetylases. The invention relates to isolated
CC rice, soybean and wheat nucleic acid fragments encoding HD1. It
CC also relates to the construction of a chimeric gene encoding all or
CC a portion of HD1, in sense or antisense orientation, where
CC expression of the chimeric gene results in production of altered
CC levels of HD1 in a transformed host cell. The availability of
CC nucleic acid sequences encoding (portions) of histone deacetylase
CC proteins will facilitate studies of global transcriptional
CC regulation in eukaryotic cells, and will also provide mechanisms to
CC control transcriptional gene regulation in plants.
XX
XX Sequence 473 AA:

```

Query Match 61.3%; Score 1558.5; DB 21; Length 473;
 Best local Similarity 61.8%; Pred. No. 1; 1e-157;
 Matches 290; Conservative 71; Mismatches 83; Indels 25; Gaps 7;

```

QY 4 DESGISLPSGPDGRKRKRSVFEYEPITIGDYYGOGHAPMKRIRIMASHLIIHNLRLKLT 63
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 esgns1pspsddvkrkksyfydrevgnuyyqgshpmkhrirmtallahy11qhmgy 61
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 SRPSLADADIGFHFSPVEVDPLASVPESMGDPSAARLRARFNNGECPVVDGJLPHPCR 123
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 lrpmaakdtdlckfhaddvva11rg1tpe1qgd--q1r1k1r1nvgedcpv1dq1ys1cq 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 124 ASAGCSIGAANKLRQDAIATNMGGGLHAKKSEASGFCYVNDIVLGLIELKMKRYL 183
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 tyagsgvgaek1knhgvcdalnmag1lhbakkceasgfcyndiv1al1el1k1her1 179
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 184 YIDIVHGDGVEAEAYTTDRVMTVSFHKFGDFPCTGHIKRVGAEKGYVALNPLNDG 243
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 yvdid1bhgdgyeaa1y1tdrvm1vaf1hk1fgdy1p1q1d1r1d1y1ak1kys11nvp1ddg 239
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 244 MDDESFRSLFRRPLCKVMEVYOEPAVVLCCGADSLGDRGFCNLSVKHACQLRPLRSY 303
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 lddesys1l1k1p1mg1kvmel1r1pgav1l1q1g1d1s1g1d1l1c1n1s1k1h1a1e1v1ym1st 299
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 304 NPLMLVGGEGYTIIRNVAKWCYETAVAVGVEDPNKLPYNEFEYFGPDYTLHVDSPME 363
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 nvp11119999ytlrnavarcwctetsvalgieddkmpqheyeyfgpdytlhvapsme 359
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 364 NLNTPKDMERIRMTLLEQLSLIHAPSVOGHTPPVNRV--LDEPEDMETRPRKPRXMSG 421
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 nksrq11delrak11dnslrlqhapsvpfqerppdterpevedgdkrwddps--- 411
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 422 TATYESDSD---DDDKPLH---GYSCRGATTPDRS--TGEDMDNDN 462
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 412 ---wdpdsdrevgdpsnpr1rvksecvd--aedkdvsgvdsma1vdep 455
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 12
AAV28800
ID AAV28800 standard; Protein: 517 AA.
XX
AC AAV28800;
XX
XX 13-JAN-2000 (first entry)
DT
DE Maize histone deacetylase-4.
XX
XX Maize histone deacetylase; HD; HD cDNA; family 1, 2mhd1; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
XX Zea mays.
XX
OS
XX
XX WO9951731-A2.
XX
PN 14-OCT-1999.
PD
PF 02-APR-1999; 99MO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
XX MPI; 1999-611038/52.
XX
DR N-PSDB; AAX90840.
XX
XX New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance
PS
PS Claim 1; Page 63-65; 87pp; English.
XX
XX The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 1, 2mhd1. This enzyme responsible for removing acetyl
CC modifications, may be localized to promoters targeted for repression by

```


